

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 15:46:05 ; Search time 54 Seconds

(without alignments)  
1465.061 Million cell updates/sec

Title: US-09-807-148-9

Perfect score: 1553

Sequence: 1 GPGTKVHVIFNYKGNVLI.....BEDKEEDBEDVPOQAKDEL 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553	100.0	280	3	AA92355
2	1553	100.0	400	3	AA92355 Recombina
3	1553	100.0	417	1	AA92355 Recombina
4	1553	100.0	417	2	AA92355 Recombina
5	1553	100.0	417	3	AA92355 Recombina
6	1553	100.0	417	5	AA92355 Recombina
7	1553	100.0	417	5	AA92355 Recombina
8	1553	100.0	417	5	AA92355 Recombina
9	1553	100.0	417	5	AA92355 Recombina
10	1553	100.0	417	6	AA92355 Recombina
11	1553	100.0	417	6	AA92355 Recombina
12	1553	100.0	417	6	AA92355 Recombina
13	1553	100.0	417	7	AA92355 Recombina
14	1553	100.0	417	7	AA92355 Recombina
15	1553	100.0	417	7	AA92355 Recombina
16	1476.5	95.1	416	7	AA92355 Recombina
17	1476.5	95.1	416	7	AA92355 Recombina
18	1442.5	92.9	401	2	AA92355 Recombina
19	1031.5	66.4	406	4	AA92355 Recombina
20	1022.5	65.8	403	2	AA92355 Recombina
21	935.5	60.2	336	4	AA92355 Recombina
22	856.5	55.2	415	4	AA92355 Recombina
23	856.5	55.2	415	4	AA92355 Recombina
24	837.5	53.9	420	5	AA92355 Recombina
25	823	53.0	312	3	AA92355 Recombina

26	823	53.0	421	3	AA92355
27	823	53.0	424	3	AA92355
28	816	52.5	312	3	AA92355
29	816	52.5	421	3	AA92355
30	816	52.5	424	3	AA92355
31	812	52.3	419	7	AA92355
32	801.5	51.6	332	3	AA92355
33	801.5	51.6	441	3	AA92355
34	801.5	51.6	444	3	AA92355
35	715	46.0	428	7	AA92355
36	702	45.2	122	2	AA92355
37	699	45.0	122	2	AA92355
38	678	43.7	122	2	AA92355
39	633	40.8	591	4	AA92355
40	629	40.5	591	7	AA92355
41	629	40.5	591	7	AA92355
42	629	40.5	591	7	AA92355
43	629	40.5	591	7	AA92355
44	629	40.5	591	7	AA92355
45	629	40.5	591	7	AA92355

## ALIGNMENTS

## RESULT 1

AA92355  
ID AAY92355 standard; protein; 280 AA.

XX  
AC AAY92355;

XX  
DT 10-AUG-2000 (first entry)

XX  
DE Recombinant delta-120 calreticulin.

XX  
KW MSP-calreticulin; maltose binding protein; angiogenesis; inhibition; endothelial cell; anti-angiogenic; neuroprotective; antidiabetic; cytosolic; dermatological; immunosuppressive; anti-inflammatory; hepatic; anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

OS Homo sapiens.

OS Synthetic.

XX  
PN WC2000020577-A1.

XX  
PD 13-APR-2000.

XX  
PF 05-OCT-1999; 99WO-US023240.

XX  
PR 06-OCT-1998; 98US-0103438P.

XX  
PA (USSH ) US-DEPT HEALTH & HUMAN SERVICES.

XX  
PI Tosato G, Pike SR, Yao L;

XX  
DR WPI; 2000-303767/26.

XX  
PT Inhibiting endothelial cell growth and angiogenesis using calreticulin, useful for suppressing tumor growth.

XX  
PS Claim 4; Page 86; 99pp; English.

XX  
CC This sequence comprises recombinant human calreticulin (AA92355) missing the N-terminal 120 amino acids. A novel method of inhibiting endothelial cell growth comprises contacting the cells with calreticulin (or its fragments/variants). Fragments of calreticulin causes at least 40% inhibition of angiogenesis, tumor growth and/or endothelial cell growth (claimed). The method may be used for inhibiting angiogenesis in a patient. The angiogenesis is associated with a disease other than a tumor that is associated with neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiodermas, immune inflammation, atherosclerosis, excessive wound repair, retinal neovascularization, macular degeneration, corneal graft

CC rejection, contact lens overwear, Crohn's disease, non-immune  
 CC inflammation, rheumatoid arthritis, systemic lupus erythematosus,  
 CC thyroiditis, Goodpasture's syndrome, systemic vasculitis, scleroderma,  
 CC Sjorgen's syndrome, sarcoidosis and primary biliary cirrhosis). The  
 CC method may also be used for treating/inhibiting tumor growth especially  
 CC Kaposi's sarcoma (claimed)  
 XX  
 SQ Sequence 280 AA;

Query Match 100.0%; Score 1553; DB 3; Length 280;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-122;  
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGTKKVVHVFYKNGKVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
 DB 1 GPGTKKVVHVFYKNGKVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60

QY 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKPEDWDEEM 120  
 DB 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKPEDWDEEM 120

QY 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWIIHPEIDNPYSPPSIYAYDNFVGLD 180  
 DB 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWIIHPEIDNPYSPPSIYAYDNFVGLD 180

QY 181 LWQVKSGTIFDNFLITNDEAYAEFFGNETGWTKAAEKQMKDKODEEORLKEEEDKKRK 240  
 DB 181 LWQVKSGTIFDNFLITNDEAYAEFFGNETGWTKAAEKQMKDKODEEORLKEEEDKKRK 240

QY 241 EEEAEDEKDEDDXDE 280  
 DB 241 EEEAEDEKDEDDXDE 280

RESULT 2  
 AAY92350  
 ID AAY92350 standard; protein; 400 AA.  
 XX  
 AC AAY92350;  
 DT 10-AUG-2000 (first entry)  
 DE Recombinant human MBP-calreticulin.  
 XX  
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;  
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;  
 KW cytostatic; dermatological; immunosuppressive; antiinflammatory, hepatic;  
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; opthalmic.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200020577-A1.  
 XX  
 PD 13-APR-2000.  
 XX  
 PF 05-OCT-1999; 99WO-US023240.  
 XX  
 PR 06-OCT-1998; 98US-0103438P.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA Tosato G, Pike SE, Yao L;  
 XX  
 PI WPI; 2000-303767/26.  
 DR N-PSDB; AAA09346, AAA09347.  
 XX  
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,  
 PT useful for suppressing tumor growth.  
 XX  
 PS Claim 4; Page 80-81; 99pp; English.  
 XX  
 CC Recombinant human MBP-calreticulin comprises the sequence of human  
 CC calreticulin (see AAY92349) without the 17 N-terminal amino acids. A

CC novel method of inhibiting endothelial cell growth comprises contacting  
 CC the cells with calreticulin (or its fragments/variants). Fragments of  
 CC calreticulin causes at least 40% inhibition of angiogenesis, tumor growth  
 CC and/or endothelial cell growth (claimed). The method may be used for  
 CC inhibiting angiogenesis in a patient. The angiogenesis is associated with  
 CC a disease other than a tumor that is associated with neovascularization  
 CC (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular  
 CC glaucoma, psoriasis, angiofibromas, immune inflammation,  
 CC atherosclerosis, excessive wound repair, retinal neovascularization,  
 CC macular degeneration, corneal graft rejection, contact lens overwear,  
 CC Crohn's disease, non-immune inflammation, rheumatoid arthritis, systemic  
 CC lupus erythematosus, thyroiditis, Goodpasture's syndrome, systemic  
 CC vasculitis, scleroderma, Sjorgen's syndrome, sarcoidosis and primary  
 CC biliary cirrhosis). The method may also be used for treating/inhibiting  
 CC tumor growth especially Kaposi's sarcoma (claimed)  
 XX  
 SQ Sequence 400 AA;

Query Match 100.0%; Score 1553; DB 3; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 4e-122;  
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGTKKVVHVFYKNGKVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
 DB 121 GPGTKKVVHVFYKNGKVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 180

QY 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKPEDWDEEM 120  
 DB 181 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKPEDWDEEM 240

QY 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWIIHPEIDNPYSPPSIYAYDNFVGLD 180  
 DB 241 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWIIHPEIDNPYSPPSIYAYDNFVGLD 300

QY 181 LWQVKSGTIFDNFLITNDEAYAEFFGNETGWTKAAEKQMKDKODEEORLKEEEDKKRK 240  
 DB 301 LWQVKSGTIFDNFLITNDEAYAEFFGNETGWTKAAEKQMKDKODEEORLKEEEDKKRK 360

QY 241 EEEAEDEKDEDDXDE 280  
 DB 361 EEEAEDEKDEDDXDE 400

RESULT 3  
 AAP92276  
 ID AAP92276 standard; protein; 417 AA.  
 XX  
 AC AAP92276;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 23-FEB-1990 (first entry)  
 XX  
 DE 60 kD Ro (Ro/SSA) antigen.  
 XX  
 KW Sjorens syndrome; systemic lupus erythematosus.  
 XX  
 OS Synthetic.  
 XX  
 PN WO8909273-A.  
 XX  
 PD 05-OCT-1989.  
 XX  
 PF 22-MAR-1989; 89WO-US001213.  
 XX  
 PR 22-MAR-1988; 88US-00171634.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Sontheimer RD, Lieu TS, Capra JD, McCauliffe DP;  
 DR WPI; 1989-309537/42.  
 DR N-PSDB; AAP92276.

PT DNA sequences encoding antigenic epitope(s) of RO 60 KD auto-antigen -  
PT used in immunoassays to detect rheumatic disease.  
XX Disclosure; Fig 2; 88pp; English.  
XX  
XX Synthetic peptides corresp. to an epitopic core of Ro antigen are  
CC expressed recombinantly to detect autoantibodies, for identification of  
CC autoimmune diseases. These epitopes are Aas 24-36, 23-36, 188-209, or 241  
CC -255. The peptides may be substd. for ribonucleoprotein particle  
CC antigens. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-  
CC MAR-2003 to correct PI field.)  
XX Sequence 417 AA;  
SQ Query Match 100.0%; Score 1553; DB 1; Length 417;  
Best Local Similarity 100.0%; Pred. No. 4.2e-122;  
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GPGTKKHVHIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 60  
DB 138 GPGTKKHVHIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 197  
QY 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDPDAKXPEDWDEEM 120  
DB 198 DDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDPDAKXPEDWDEEM 257  
QY 121 DGEWPEPVIQNPYKGEWKPRQIDNPDKYKGTWIHPEIDNPESDPDSIYAYDNFVGLGLD 180  
DB 258 DGEWPEPVIQNPYKGEWKPRQIDNPDKYKGTWIHPEIDNPESDPDSIYAYDNFVGLGLD 317  
QY 181 LMQVKSGTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKDKODEQRLEKEEEDKKRK 240  
DB 318 LMQVKSGTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKDKODEQRLEKEEEDKKRK 377  
RESULT 4  
AAAY00927  
ID AAAY00927 standard; protein; 417 AA.  
XX  
XX AAAY00927;  
AC Calreticulin.  
XX  
XX C1q and collectin receptor; cClqR binding domain; complement ubiquitin;  
XX CUB functionality; inhibitor; complement activation; inflammation;  
XX myocardial infarction; brain ischaemia; gut ischaemia; amyloid plaque;  
XX rheumatoid arthritis; systemic lupus erythematosus; Alzheimer's disease;  
XX immune complex nephritis; therapy.  
XX Homo sapiens.  
XX WO9507406-A1.  
XX 18-FEB-1999.  
XX  
XX 12-AUG-1999; 98WO-GB002430.  
XX  
XX 12-AUG-1997; 97GB-00016998.  
XX  
XX (UYLE-) UNIV LEICESTER.  
XX  
XX Schwaeble W;  
XX  
XX WPI; 1999-180404/15.  
XX  
XX Use of a cClqR binding domain - to modulate complement ubiquitin (CUB)  
PT functionality.

XX Disclosure; Page 26-27; 31pp; English.  
XX  
XX This sequence is calreticulin, a homologue of Clq and collectin receptor  
CC (cClqR). The invention relates to the use of a cClqR binding domain in a  
CC medicament to effect complement ubiquitin (CUB) functionality, and an  
CC inhibitor of the cClqR binding domain in a medicament to inhibit CUB  
CC functionality. The cClqR binding domain, or its inhibitor, can be used to  
CC treat a human or animal body. Particularly an inhibitor is used to treat  
CC complement activation involved in the initiation and maintenance of  
CC inflammation, for example in myocardial infarction, brain ischaemia  
CC (stroke), gut ischaemia, rheumatoid arthritis, systemic lupus  
CC erythematosus, burns, immune complex nephritis, and to treat amyloid  
CC plaques in Alzheimer's disease. The use of cClqR binding domain or  
CC inhibitor enables the CUB domain functionality to be modulated using a  
CC low molecular weight molecule  
XX  
XX Sequence 417 AA;  
SQ Query Match 100.0%; Score 1553; DB 2; Length 417;  
Best Local Similarity 100.0%; Pred. No. 4.2e-122;  
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GPGTKKHVHIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 60  
DB 138 GPGTKKHVHIFNYKGNVLINKDIRCKDDEFTHLYTLIVRPDNTYEVKIDNSQVSGSLE 197  
QY 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDPDAKXPEDWDEEM 120  
DB 198 DDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWPKPEHIPPDPDAKXPEDWDEEM 257  
QY 121 DGEWPEPVIQNPYKGEWKPRQIDNPDKYKGTWIHPEIDNPESDPDSIYAYDNFVGLGLD 180  
DB 258 DGEWPEPVIQNPYKGEWKPRQIDNPDKYKGTWIHPEIDNPESDPDSIYAYDNFVGLGLD 317  
QY 181 LMQVKSGTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKDKODEQRLEKEEEDKKRK 240  
DB 318 LMQVKSGTIFDNFLITNDEAYAEFGNETGWTKAAEKQMKDKODEQRLEKEEEDKKRK 377  
RESULT 5  
AAAY92349  
ID AAAY92349 standard; protein; 417 AA.  
XX  
XX AAAY92349;  
AC  
XX 10-AUG-2000 (first entry)  
XX Human MBP-calreticulin.  
XX  
XX MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;  
XX endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;  
XX cyrostatic; dermatological; immunosuppressive; antiinflammatory; hepatic;  
XX anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.  
XX  
XX Homo sapiens.  
XX  
XX Location/Qualifiers  
PH Key 1..17  
FT Peptide /label= signal\_peptide  
FT Protein 18  
FT /label= mature\_protein  
XX  
XX WO200020577-A1.  
XX  
XX 13-APR-2000.  
XX  
XX 05-OCT-1999; 99WO-US023240.  
XX  
XX

PR 06-OCT-1998; 98US-0103438P.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Tosato G, Pike SE, Yao L;  
 XX  
 DR WPI; 2000-303767/26.  
 DR N-PSDB; AAA09346, AAA09347.  
 XX  
 XX Inhibiting endothelial cell growth and angiogenesis using calreticulin,  
 PT useful for suppressing tumor growth.  
 XX  
 PS Disclosure; Page 79-80; 99pp; English.  
 XX  
 CC A novel method of inhibiting endothelial cell growth comprises contacting  
 CC the cells with calreticulin (or its fragments/variants). Fragments of  
 CC calreticulin causes at least 40% inhibition of angiogenesis, tumor growth  
 CC and/or endothelial cell growth (claimed). The method may be used for  
 CC inhibiting angiogenesis in a patient. The angiogenesis is associated with  
 CC a disease other than a tumor that is associated with neovascularization  
 CC (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular  
 CC glaucoma, psoriasis, angiofibromas, immune inflammation,  
 CC atherosclerosis, excessive wound repair, retinal neovascularization,  
 CC macular degeneration, corneal graft rejection, contact lens overwear,  
 CC Crohn's disease, non-immune inflammation, rheumatoid arthritis, systemic  
 CC lupus erythematosus, thyroiditis, Goodpasture's syndrome, systemic  
 CC vasculitis, scleroderma, Sjorgen's syndrome, sarcoidosis and primary  
 CC biliary cirrhosis). The method may also be used for treating/inhibiting  
 CC tumor growth especially Kaposi's sarcoma (claimed)  
 XX  
 SQ Sequence 417 AA;  
 Query Match 100.0%; Score 1553; DB 3; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-122;  
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GPGTKKHVIFNYKGNVLINKDKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSLE 60  
 DB 138 GPGTKKHVIFNYKGNVLINKDKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSLE 197  
 QY 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDPKPEHIPDPDAKPEDWDEEM 120  
 DB 198 DDWDFLPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDPKPEHIPDPDAKPEDWDEEM 257  
 QY 121 DGEWEPPIQNPYKGEWKPRQIDNPYKGTWTHPEIDNPYSPDPSIYADNFGVLGLD 180  
 DB 258 DGEWEPPIQNPYKGEWKPRQIDNPYKGTWTHPEIDNPYSPDPSIYADNFGVLGLD 317  
 QY 181 LWQVKSgtIFDNFLITNDDEAYAEFGNETWGTVAEKQMKDKQDEBQRLKEEEDKKRK 240  
 DB 318 LWQVKSgtIFDNFLITNDDEAYAEFGNETWGTVAEKQMKDKQDEBQRLKEEEDKKRK 377  
 QY 241 EEEAEADKEDDEDKDEEEDKEEEDVPGQAKDEL 280  
 DB 378 EEEAEADKEDDEDKDEEEDKEEEDVPGQAKDEL 417  
 RESULT 6  
 AAU77712  
 ID AAU77712 standard; protein; 417 AA.  
 XX  
 AC AAU77712;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Human calreticulin (CRT).  
 XX  
 KW Calreticulin; CRT; endoplasmic reticulum chaperone polypeptide;  
 KW cytosolic; vaccine; human papillomavirus 16; HPV 16; E7; DNA vaccine;  
 KW enhanced antigen-specific immune response; cytotoxic T lymphocyte;  
 KW tumour; cancer; cervical cancer.  
 XX  
 OS Homo sapiens.

XX WC200212281-A2.  
 PN  
 XX  
 PD 14-FEB-2002.  
 XX  
 XX 02-AUG-2001; 2001WO-US024134.  
 XX PF  
 XX 03-AUG-2000; 2000US-0222902P.  
 XX PR  
 XX (UYJO ) UNIV JOHNS HOPKINS.  
 PA  
 XX  
 PI Wu T, Hung C;  
 XX  
 XX WPI; 2002-257463/30.  
 DR N-PSDB; ABK11662.  
 XX  
 XX New nucleic acids encoding a fusion polypeptide comprising an endoplasmic  
 PT reticulum chaperone polypeptide linked to an antigenic polypeptide,  
 PT useful as a vaccine for inducing antigen-specific immune responses.  
 XX  
 PS Disclosure; Page 27; 71pp; English.  
 XX  
 CC The invention describes a nucleic acid molecule (I) encoding a fusion  
 CC polypeptide comprising a first polypeptide domain comprising an  
 CC endoplasmic reticulum chaperone polypeptide e.g. calreticulin (CRT) and a  
 CC second polypeptide domain comprising at least one antigenic peptide e.g.  
 CC Human papillomavirus 16 (HPV 16) E7. The nucleic acid is useful as a  
 CC vaccine (DNA vaccine) for inducing enhanced antigen-specific immune  
 CC responses, particularly those mediated by cytotoxic T lymphocytes. The  
 CC nucleic acid and compositions comprising the nucleic acid is also useful  
 CC for inhibiting the growth of tumours and cancers e.g. cervical cancer.  
 CC This is the amino acid sequence of the human calreticulin (CRT), an  
 CC endoplasmic reticulum protein used in the creation of a DNA vaccine  
 XX  
 SQ Sequence 417 AA;  
 Query Match 100.0%; Score 1553; DB 5; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-122;  
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GPGTKKHVIFNYKGNVLINKDKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSLE 60  
 DB 138 GPGTKKHVIFNYKGNVLINKDKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSLE 197  
 QY 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDPKPEHIPDPDAKPEDWDEEM 120  
 DB 198 DDWDFLPPKKIKDPDASKPEDWDERAKIDPTDTSKPEDWDPKPEHIPDPDAKPEDWDEEM 257  
 QY 121 DGEWEPPIQNPYKGEWKPRQIDNPYKGTWTHPEIDNPYSPDPSIYADNFGVLGLD 180  
 DB 258 DGEWEPPIQNPYKGEWKPRQIDNPYKGTWTHPEIDNPYSPDPSIYADNFGVLGLD 317  
 QY 181 LWQVKSgtIFDNFLITNDDEAYAEFGNETWGTVAEKQMKDKQDEBQRLKEEEDKKRK 240  
 DB 318 LWQVKSgtIFDNFLITNDDEAYAEFGNETWGTVAEKQMKDKQDEBQRLKEEEDKKRK 377  
 QY 241 EEEAEADKEDDEDKDEEEDKEEEDVPGQAKDEL 280  
 DB 378 EEEAEADKEDDEDKDEEEDKEEEDVPGQAKDEL 417  
 RESULT 7  
 AAU77712  
 ID AAU77712 standard; protein; 417 AA.  
 XX  
 AC AAU77712;  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human calreticulin protein.  
 XX  
 KW Human; calreticulin; antisense compound; hyperproliferative disorder;  
 KW cancer; autoimmune disease; viral infection; cardiovascular disease;





CC diagnosis, therapy and pharmaceutical development. The modulators are  
CC useful in a variety of diagnostic and therapeutic applications including  
CC angiogenic, apoptotic and cell proliferation disorders. This sequence  
CC represents an MP21 protein of the invention  
XX  
SQ Sequence 417 AA;  
  
Query Match 100.0%; Score 1553; DB 6; Length 417;  
Best Local Similarity 100.0%; Pred. No. 4.2e-122;  
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GPGTKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60  
Db 138 GPGTKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 197  
  
Qy 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDPKEHIPDPDAKPEDWDEEM 120  
Db 198 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDPKEHIPDPDAKPEDWDEEM 257  
  
Qy 121 DGEWPEPPVIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPESDPDSIYAYDNFVGLGLD 180  
Db 258 DGEWPEPPVIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPESDPDSIYAYDNFVGLGLD 317  
  
Qy 181 LWQVKSGTIFDNFLITNDAYAEFGNETWGTAKAEKQMKQKQDEEQLKEEEDKKRK 240  
Db 318 LWQVKSGTIFDNFLITNDAYAEFGNETWGTAKAEKQMKQKQDEEQLKEEEDKKRK 377  
  
Qy 241 EEEAEKDEDDKDEDEEDKDEEEDKDEEEDVPGQAKDEL 280  
Db 378 EEEAEKDEDDKDEDEEDKDEEEDKDEEEDVPGQAKDEL 417  
  
RESULT 11  
ID AAG79824 standard; protein; 417 AA.  
XX  
AC AAG79824;  
XX  
DT 16-APR-2003 (first entry)  
XX  
DE Calreticulin.  
XX  
KW Motif: SE; signal transduction; Alzheimer's disease; neuronal cell;  
KW calreticulin; arrhythmia; idiopathic nephritic syndrome;  
KW non-autoimmune hyperthyroidism; obesity; polycystic kidney disease;  
KW cancer; asthma; atopic dermatitis; psoriasis; rheumatoid arthritis;  
KW systemic lupus erythematosus; infectious disease; atherosclerosis;  
KW inflammatory bowel disease; osteoarthritis; septic shock;  
KW congestive heart failure; insulin-resistance syndrome;  
KW ischaemia-reperfusion injury.  
XX  
OS Homo sapiens.  
XX  
FN WO200299061-A2.  
XX  
PD 12-DEC-2002.  
XX  
PF 04-JUN-2002; 2002WO-US017536.  
XX  
PR 04-JUN-2001; 2001US-0295691P.  
XX  
PR 03-JUN-2002; 2002US-00161959.  
XX  
PA (UNWI ) UNIV MICHIGAN.  
XX  
PI Holoshitz J, Ling S;  
XX  
DR WPI; 2003-156953/15.  
XX  
PT Treating diseases associated with signal transduction aberrations, e.g.  
PT Alzheimer's disease, arrhythmia or rheumatoid arthritis, comprises  
PT administering a preparation comprising an SE- or an SE motif-containing  
PT peptide.  
XX

PS Disclosure; Fig 14; 97pp; English.  
XX  
CC This sequence represents human calreticulin. This protein was used in the  
CC method of the invention for treating diseases associated with signal  
CC transduction aberrations. The method comprises: (a) providing a subject  
CC with one or more symptoms of Alzheimer's disease and, optionally, a  
CC plurality of neuronal cells expressing calreticulin, and a preparation  
CC comprising an shared epitope (SE) - containing peptide or a peptide which  
CC binds the calreticulin; and (b) administering the preparation to the  
CC subject. The inventive method is useful for treating diseases associated  
CC with signal transduction aberrations, such as Alzheimer's disease,  
CC arrhythmia, idiopathic nephritic syndrome, non-autoimmune  
CC hyperthyroidism, obesity, polycystic kidney disease, cancer, asthma,  
CC atopic dermatitis, psoriasis, rheumatoid arthritis, systemic lupus  
CC erythematosus, infectious diseases, inflammatory bowel disease,  
CC osteoarthritis, septic shock, atherosclerosis, congestive heart failure,  
CC insulin-resistance syndrome, and ischaemia-reperfusion injury  
XX  
SQ Sequence 417 AA;  
  
Query Match 100.0%; Score 1553; DB 6; Length 417;  
Best Local Similarity 100.0%; Pred. No. 4.2e-122;  
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GPGTKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 60  
Db 138 GPGTKVHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE 197  
  
Qy 61 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDPKEHIPDPDAKPEDWDEEM 120  
Db 198 DDWDFLPPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDPKEHIPDPDAKPEDWDEEM 257  
  
Qy 121 DGEWPEPPVIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPESDPDSIYAYDNFVGLGLD 180  
Db 258 DGEWPEPPVIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPESDPDSIYAYDNFVGLGLD 317  
  
Qy 181 LWQVKSGTIFDNFLITNDAYAEFGNETWGTAKAEKQMKQKQDEEQLKEEEDKKRK 240  
Db 318 LWQVKSGTIFDNFLITNDAYAEFGNETWGTAKAEKQMKQKQDEEQLKEEEDKKRK 377  
  
Qy 241 EEEAEKDEDDKDEDEEDKDEEEDKDEEEDVPGQAKDEL 280  
Db 378 EEEAEKDEDDKDEDEEDKDEEEDKDEEEDVPGQAKDEL 417  
  
RESULT 12  
ID ADA26337 standard; protein; 417 AA.  
XX  
AC ADA26337;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human calreticulin protein SEQ ID NO:14.  
XX  
KW stress response protein; immune response; tumour;  
KW tumour growth inhibition; metastasis; cytostatic; vaccine; gene therapy;  
KW human; calreticulin.  
XX  
OS Homo sapiens.  
XX  
FN WO2003068941-A2.  
XX  
PD 21-AUG-2003.  
XX  
PF 13-FEB-2003; 2003WO-US004631.  
XX  
PR 13-FEB-2002; 2002US-0356293P.  
XX  
PA (UYDU-) UNIV DUKE.  
XX  
PI Nicchita C, Baker-Lepain J;  
XX

DR WPI; 2003-679635/64.  
 DR N-PSDB; ADA26336.  
 XX  
 PT New Hsp60, Hsp70, Hsp90 or calreticulin polypeptide, useful for preparing  
 PT a composition for inhibiting tumor growth or metastasis in a subject.  
 XX  
 PS Disclosure; Page 137-138; 150pp; English.  
 XX  
 CC The present invention describes a recombinant stress response polypeptide  
 CC (I) free of an antigen-binding domain, where the recombinant stress  
 CC response polypeptide comprises an extracellularly transported polypeptide  
 CC when expressed in a host cell. Also described: (1) a composition for  
 CC eliciting an immune response in a subject; (2) eliciting an immune  
 CC response in a subject; and (3) inhibiting tumour growth or metastasis in  
 CC a subject. (1) has cytostatic activity, and can be used in vaccines and  
 CC in gene therapy. The polypeptide (I) can be used for preparing a  
 CC composition for inhibiting tumour growth or metastasis in a subject. The  
 CC present sequence represents human calreticulin, which is a stress  
 CC response polypeptide given in the exemplification of the present  
 CC invention.  
 XX  
 SQ Sequence 417 AA;  
 Query Match 100.0%; Score 1553; DB 6; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-122;  
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GPGTKKHVIFNYKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
 DB 138 GPGTKKHVIFNYKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 197  
 QY 61 DDWDFLPFKIKDPDASKPEDWDERAKIDDDTDSKPEDWDKPEHIPDPDAKPEDWDEEM 120  
 DB 198 DDWDFLPFKIKDPDASKPEDWDERAKIDDDTDSKPEDWDKPEHIPDPDAKPEDWDEEM 257  
 QY 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWIIHPEIDNPEYSPDPSIAYDNFGVLGLD 180  
 DB 318 LWQVKSGETIFNFLTNDDEAYAEFGNETGWTKAAEKQMKDKQDEBQRLKEEEDKKRK 377  
 QY 241 EEEAEADKEDDEDKDEDEEDKEDEEEDVPGQAKDEL 280  
 DB 378 EEEAEADKEDDEDKDEDEEDKEDEEEDVPGQAKDEL 417  
 RESULT 13  
 ADD22407  
 ID ADD22407 standard; protein; 417 AA.  
 AC  
 AC ADD22407;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE HLA-B46 T cell recognised tumour antigenic polypeptide, SEQ No 57.  
 XX  
 KW tumour antigenic peptide; cancer; vaccine; cytostatic; cytotoxic T cell;  
 KW colon; mouth; lung; prostatic; gynecological; human.  
 XX  
 OS Homo sapiens.  
 XX  
 FN JP2003111595-A.  
 XX  
 PD 15-APR-2003.  
 XX  
 PF 24-JUN-2002; 2002JP-00183603.  
 XX  
 PR 25-JUN-2001; 2001JP-00191974.  
 XX  
 FA (ITCY/) ITO Y.  
 XX

DR WPI; 2003-611129/58.  
 XX  
 PT Novel tumor antigenic peptide or polypeptide useful for inducing  
 PT cytotoxic T cells or for treating cancer such as colon, mouth, lung,  
 PT prostatic or gynecological cancer.  
 XX  
 PS Claim 2; SEQ ID NO 57; 98pp; Japanese.  
 XX  
 CC The invention relates to a novel tumour antigenic peptide or polypeptide  
 CC comprising a sequence selected from 99 sequences fully defined in the  
 CC specification. The tumour antigenic peptide or polypeptide comprises a  
 CC sequence selected from 99 sequences fully defined in the specification,  
 CC where the tumour antigenic peptide preferably has a sequence of Glu-Pro-  
 CC pro-Leu-Ser-Gln-Glu-Thr-Phe, and the polypeptide preferably has a  
 CC sequence comprising 393 amino acids fully defined in the specification.  
 CC The invention further provides a cancer vaccine comprising a tumour  
 CC antigenic peptide or polypeptide, which has cytostatic activity. A tumour  
 CC antigenic peptide, polypeptide, its encoding polynucleotide, a  
 CC hybridising polynucleotide, a recombinant vector containing the  
 CC polynucleotide, a host transformed with the vector or an antibody are  
 CC useful for screening for compounds that interact with the tumour  
 CC antigenic peptide, the polypeptide or its encoding polynucleotide and  
 CC increases the expression of the tumour antigenic peptide, the polypeptide  
 CC or polynucleotide. The tumour antigenic peptide or the polypeptide is  
 CC useful for inducing cytotoxic T cells. The tumour antigenic peptide  
 CC vaccine is useful for treating cancer such as colon, mouth, lung,  
 CC prostatic or gynecological cancer. The invention also provides a  
 CC pharmaceutical composition useful for treating cancer. The tumour  
 CC antigenic peptide or the polypeptide is useful as an antigen to create  
 CC antibodies. This sequence represents one of the tumour antigenic  
 CC polypeptides of the invention.  
 XX  
 SQ Sequence 417 AA;  
 Query Match 100.0%; Score 1553; DB 7; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-122; Indels 0; Gaps 0;  
 Matches 280; Conservative 0; Mismatches 0;  
 QY 1 GPGTKKHVIFNYKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
 DB 138 GPGTKKHVIFNYKGNVLINKIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 197  
 QY 61 DDWDFLPFKIKDPDASKPEDWDERAKIDDDTDSKPEDWDKPEHIPDPDAKPEDWDEEM 120  
 DB 198 DDWDFLPFKIKDPDASKPEDWDERAKIDDDTDSKPEDWDKPEHIPDPDAKPEDWDEEM 257  
 QY 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWIIHPEIDNPEYSPDPSIAYDNFGVLGLD 180  
 DB 258 DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWIIHPEIDNPEYSPDPSIAYDNFGVLGLD 317  
 QY 181 LMQVKSGETIFNFLTNDDEAYAEFGNETGWTKAAEKQMKDKQDEBQRLKEEEDKKRK 240  
 DB 318 LMQVKSGETIFNFLTNDDEAYAEFGNETGWTKAAEKQMKDKQDEBQRLKEEEDKKRK 377  
 QY 241 EEEAEADKEDDEDKDEDEEDKEDEEEDVPGQAKDEL 280  
 DB 378 EEEAEADKEDDEDKDEDEEDKEDEEEDVPGQAKDEL 417  
 RESULT 14  
 ADE56308  
 ID ADE56308 standard; protein; 417 AA.  
 XX  
 AC ADE56308;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein NP\_004334, SEQ ID NO 2160.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX





```
XX
SQ Sequence 417 AA;
Query Match      100.0%; Score 1553; DB 7; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.2e-122;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
Db 138 GPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSLE 197

Qy 61 DDWDFLPPKKIKDPASKPEDWDERAKIDDDPTDSKPEDWPKPEHIPDPDAKPEDWDEEM 120
Db 198 DDWDFLPPKKIKDPASKPEDWDERAKIDDDPTDSKPEDWPKPEHIPDPDAKPEDWDEEM 257

Qy 121 DGEWEPPIQNPYKGEWKPRQIDNPDYKGTWIIHPEIDNPEYSPDPSIYAYDNFGVLGLD 180
Db 258 DGEWEPPIQNPYKGEWKPRQIDNPDYKGTWIIHPEIDNPEYSPDPSIYAYDNFGVLGLD 317

Qy 181 LWQVKSGTIFDNFLITNDEAYAEFFGNETWGTVAEAKQMKDKODEEQRLEKEEEDKKRK 240
Db 318 LWQVKSGTIFDNFLITNDEAYAEFFGNETWGTVAEAKQMKDKODEEQRLEKEEEDKKRK 377

Qy 241 EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
Db 378 EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 417
```

Search completed: October 12, 2004, 15:48:29  
Job time : 56 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 15:46:06 ; Search time 19 Seconds  
(without alignments)  
760.804 Million cell updates/sec

Title: US-09-807-148-9  
Perfect score: 1553  
Sequence: 1 GPCTKKVHVFNYKGNVLI.....EEDKEDEBEDVFGQAKDEL 280

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pap.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pap.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pap.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pap.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pap.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1549	99.7	416	US-09-828-000-2	Sequence 2, Appli
2	856.5	55.2	415	US-08-675-816-2	Sequence 2, Appli
3	624	40.2	593	US-08-296-382-2	Sequence 2, Appli
4	577.5	37.2	610	US-09-976-594-947	Sequence 947, App
5	493	31.7	542	US-08-675-816-6	Sequence 6, Appli
6	318	20.5	61	US-09-828-000-8	Sequence 8, Appli
7	318	20.5	180	US-09-828-000-3	Sequence 3, Appli
8	235	15.1	61	US-09-828-000-4	Sequence 4, Appli
9	189	12.2	35	US-09-828-000-7	Sequence 7, Appli
10	184.5	11.9	714	US-08-990-114-3	Sequence 3, Appli
11	184.5	11.9	714	US-09-241-333-3	Sequence 3, Appli
12	184.5	11.9	740	US-08-257-073-5	Sequence 5, Appli
13	175.5	11.3	1162	US-08-728-323A-2	Sequence 2, Appli
14	175.5	11.3	1162	US-09-298-568-2	Sequence 2, Appli
15	175.5	11.3	1162	US-09-410-395-2	Sequence 2, Appli
16	163	10.5	905	US-08-574-959A-9	Sequence 9, Appli
17	163	10.5	905	US-09-357-014-9	Sequence 9, Appli
18	163	10.5	1135	US-08-574-959A-7	Sequence 7, Appli
19	163	10.5	1135	US-09-357-014-7	Sequence 7, Appli
20	151	9.7	740	US-09-022-983-5	Sequence 5, Appli
21	150.5	9.7	546	US-08-935-855-20	Sequence 20, Appl
22	149	9.6	411	US-08-741-134-6	Sequence 6, Appli
23	146	9.4	754	US-09-214-564A-2	Sequence 2, Appli
24	146	9.4	764	US-09-370-803-67	Sequence 67, Appl
25	143.5	9.2	197	US-09-486-147-40	Sequence 40, Appl
26	143.5	9.2	240	US-08-114-555A-8	Sequence 8, Appli
27	143.5	9.2	240	US-08-559-397A-14	Sequence 14, Appl

28	142	9.1	27	4	US-09-828-000-5	Sequence 5, Appli
29	137	8.8	633	1	US-08-458-477A-5	Sequence 5, Appli
30	137	8.8	633	2	US-09-033-153-5	Sequence 5, Appli
31	137	8.8	633	3	US-09-325-430B-5	Sequence 5, Appli
32	137	8.8	765	2	US-08-663-112-2	Sequence 2, Appli
33	135	8.7	295	2	US-08-679-765-5	Sequence 5, Appli
34	135	8.7	295	3	US-09-196-525-5	Sequence 5, Appli
35	135	8.7	295	3	US-09-318-317-5	Sequence 5, Appli
36	135	8.7	295	4	US-09-177-165A-22	Sequence 22, Appl
37	134.5	8.7	231	3	US-09-461-697-194	Sequence 194, App
38	134.5	8.7	232	3	US-09-461-697-192	Sequence 192, App
39	134.5	8.7	238	3	US-09-461-697-190	Sequence 190, App
40	134.5	8.7	257	3	US-09-461-697-188	Sequence 188, App
41	134.5	8.7	272	3	US-09-461-697-186	Sequence 186, App
42	133.5	8.6	1312	4	US-09-345-882-29	Sequence 29, Appl
43	133	8.6	567	4	US-09-205-258-573	Sequence 573, App
44	132.5	8.5	558	1	US-08-285-440-6	Sequence 6, Appli
45	132.5	8.5	558	1	US-08-630-343-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-09-828-000-2  
; Sequence 2, Application US/09828000  
; Patent No. 6596690  
; GENERAL INFORMATION:  
; APPLICANT: Government of the United States of America  
; TITLE OF INVENTION: Vasostatin as Marrow Protectant  
; FILE REFERENCE: 4239-55414  
; CURRENT APPLICATION NUMBER: US/09/828,000  
; CURRENT FILING DATE: 2001-04-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Calreticulin  
US-09-828-000-2

Query Match 99.7%; Score 1549; DB 4; Length 416;  
Best Local Similarity 100.0%; Pred. No. 4.4e-130;  
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GPCTKKVHVFNYKGNVLI	IRCKDDEFTHTLYTLIVR	PDNTYEVKIDNSQVSGSLE	60
Db	138	GPCTKKVHVFNYKGNVLI	IRCKDDEFTHTLYTLIVR	PDNTYEVKIDNSQVSGSLE	197
Qy	61	DDWDELPPKKIKDPDASK	PDWDERAKIDPDTDSKPED	WDKPEHI	120
Db	198	DDWDELPPKKIKDPDASK	PDWDERAKIDPDTDSKPED	WDKPEHI	257
Qy	121	DGEWEPPIQNPEYKGEWK	PRQIDNPYKGTWTHPEID	NPESPDPSIAYDNFVGLD	180
Db	258	DGEWEPPIQNPEYKGEWK	PRQIDNPYKGTWTHPEID	NPESPDPSIAYDNFVGLD	317
Qy	181	LWQVKSFTFDNLIINDEA	YAEFGNETWGTVAEKQK	MDKQDEEQR	240
Db	318	LWQVKSFTFDNLIINDEA	YAEFGNETWGTVAEKQK	MDKQDEEQR	377
Qy	241	EEEEAEKDEDEDEDEDE	DEDEDEDEDEDEDEDE	DEDEDEDEDEDEDEDE	279
Db	378	EEEEAEKDEDEDEDEDE	DEDEDEDEDEDEDEDE	DEDEDEDEDEDEDEDE	416

RESULT 2  
US-08-675-816-2  
; Sequence 2, Application US/08675816  
; Patent No. 6171864  
; GENERAL INFORMATION:  
; APPLICANT: Coughlan, Sean J.  
; APPLICANT: Winfrey, Jr., Ron J.

;; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND USES

;; NUMBER OF SEQUENCES: 15  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Seed and Berry  
;; STREET: 701 Fifth Ave. Suite 6300  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: U.S.A.  
;; ZIP: 98104-7092  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/675,816  
;; FILING DATE: 05-JUL-1996  
;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 6171864tenburg, Carol  
;; REGISTRATION NUMBER: 39,317  
;; REFERENCE/DOCKET NUMBER: 750027.401  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206)-622-4900  
;; TELEFAX: (206)-682-6031  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 415 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-675-816-2

Query Match 55.2%; Score 856.5; DB 3; Length 415;

Best Local Similarity 55.3%; Pred. No. 2.2e-68;

Matches 157; Conservative 36; Mismatches 80; Indels 11; Gaps 4;

QY 1 GPGTKKHVIVNYGKGNVLINIKRDKDDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSL 60

DB 139 GYSTKVKHAILNYNDTHLIKVEPCETDQTHVYTLVIRPDNTYEVKIDNSQVSGSL 198

QY 61 DDWFLPKKIDPDASKPEDWDERAKIDDPDTSKPEDWPK-PEHIDPDPAKPEDWDER 119

DB 199 TDWLLPPLPKIDPEAKKPDWDEKEIYDPEDKKPGYDDIPEKIDPDPAKPEDWDE 258

QY 120 MDGEWPEPVIQNPYKGEWPKQIDNPYKGTWHPIDNPYSPPFSIYAYDNFVGLG 179

DB 259 EDGEWTAPTIANPEYKGPWKPKIKPNYKKGWKAPMIDNPDPKDDPEIYVYVNLKYVGI 318

QY 180 DLWQVKSFTIDNFIITNDAYAEFGNETGWTKAAEKQMKQKQDEQELKEEEDKKR 239

DB 319 ELWQVKSFTIDNFIITNDAYAEFGNETGWTKAAEKQMKQKQDEQELKEEEDKKR 375

QY 240 KEEBEEADKEDDEKDEDEDEDEKE---EDEEDVPGQAKDEL 280

DB 376 PADSDADDDDDDTSGEDGSGSDAEDSAEDV-----HDEL 415

## RESULT 3

US-08-296-362-2

Sequence 2, Application US/08296362

Patent No. 5691306

GENERAL INFORMATION:

APPLICANT: Bergeron, John J.M.

APPLICANT: Thomas, David Y.

APPLICANT: Waga, Ikuo

TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF

TITLE OF INVENTION: PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY

TITLE OF INVENTION: PROTEIN PRODUCTION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY

STREET: 6300 Columbia Center, 701 Fifth Avenue

;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: US  
;; ZIP: 98104-7092  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/296,362  
;; FILING DATE: 25-AUG-1994  
;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Deehr, Yanya S.

;; REGISTRATION NUMBER: 37,120

;; REFERENCE/DOCKET NUMBER: 690066.401C1

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (206) 622-4900

;; TELEFAX: (206) 682-6031

;; TELEX: 3723836

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 593 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

;; US-08-296-362-2

Query Match 40.2%; Score 624; DB 1; Length 593;

Best Local Similarity 33.7%; Pred. No. 1.8e-47;

Matches 129; Conservative 58; Mismatches 86; Indels 110; Gaps 8;

QY 6 KVHVFNYGKGNVLINIKRDK-----DDEFTHTLYTLVIRPDNTYEVKIDNSQV 55

DB 200 KUHFIRHNPKTGYEEKHAKRPDADLKTYFTKKHTLYTLINPDNFSFELVDQSLN 259

QY 56 SGLSEDDWDFPP-----KKIKDPDASKPEDWDERAKIDDPDTSKPEDW----- 99

DB 260 SGNLLN--DMTPVPNPSREIDEPDQKPEDWDERPKIPDPDAVKPDWEDAPAKIPDEE 317

QY 100 -----DKPEHIDPDPAKPEDWDEMGWEP-----PVIQ 131

DB 318 ATKPGWLLDDPEYVYVDPDAEKPEDWEDMDGWEAPOIANPKCESAPGCGVWQRPMDN 377

QY 132 EYKGEWPKQIDNPYKGTWHPIDNPYSPPFSIYAYDNFVGLGLDQVKSFTID 191

DB 378 PNYKKGWKPMIDNPYQGIWKPKIPNPDFFEDLEPFKMTFFSAIGLEWMTSDIFFD 437

QY 192 NPLITNDAYAEFGNETGWTKAAE-----KQMKDQDE----- 226

DB 438 NFIVCGRRVVDWANDGWLKKAADGAAPGVVGQMIETAAEERPWLWVYVLTVALPVF 497

QY 227 -----EQLKEEEDK-----KRKEEBAEDKEDDEDDED 257

DB 498 LVISFCCSGKQSPVEYKKTDAQPDPVKEEKEEKEKOKGDEEESGEKLEBKQKSDA 557

## RESULT 4

US-09-976-594-947

Sequence 947, Application US/09976594

Patent No. 6673549

GENERAL INFORMATION:

APPLICANT: Furness, Michael

APPLICANT: Buchbinder, Jenny

TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY

STREET: 6300 Columbia Center, 701 Fifth Avenue

REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 750027.401  
TELEPHONE: (206)-622-4900  
TELEFAX: (206)-682-6031  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 542 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-675-816-6

Query Match 31.7%; Score 493; DB 3; Length 542;  
Best Local Similarity 29.8%; Pred. No. 7.6e-36;  
Matches 117; Conservative 41; Mismatches 113; Indels 122; Gaps 10;

QY 1 GP-----GKVKVHVFNYKG--KNVLINKDIR-----CKODEFTHLYTLIVRPDNTYEVKID 50  
DB 143 GPDKCGATNKVHFKHKNPKSGEVIHHLKYPPSPDKLTHVYTAILKPDNEURLVD 202

QY 51 NSQVESGSLEDDWDLP-----KTIKDPDASKPEDWDERAKIDDPDTSKPEDW----- 99  
DB 203 GEEKKANFLSSDDPEPLVPKTIPTDDKKPEDWDERAKIPDPNAVKPDDEADAPME 262

QY 100 -----DKPEHIPDPDAKKPEDWDEMDGEWEPVIONPE----- 133  
DB 263 IVDEDAEKPEGWLDDEPEBIDDDAAKPEDWDEMDGEWEPKIDNPKCETAPGCGEWR 322

QY 134 -----YGEWKPRIDNDPKYKGTWHPHIDNPEY-----SPDPSIYVDNFGVLGLDLW 182  
DB 323 PMCRNPAYKKGKHPALIDNPYKGIWKPOEIPNPNTFELEKPD-----PEPIAAVGIEIW 377

QY 183 QVKSQGITFONFLITNDEAYAEFGNETGWVTKAAEKQMKDKQD----- 225  
DB 378 TMDGGLIFGHILLADDEKVAESLRQTAWKPKFDAEKEKQAEADAAGSDLAGFKKQVFD 437

QY 226 -----BEQRKEEEDK----- 243  
DB 438 LLYQVADIPFLSHDKIIDIIEKGEKQPNLTIGILVIVVFTVLFKILFGGKKPAKV 497

RESULT 6  
US-09-828-000-8  
Sequence 8, Application US/09828000  
Patent No. 6596690  
GENERAL INFORMATION:  
APPLICANT: Government of the United States of America  
TITLE OF INVENTION: Vasostatin as Marrow Protectant  
FILE REFERENCE: 4239-55414  
CURRENT APPLICATION NUMBER: US/09/828,000  
CURRENT FILING DATE: 2001-04-06  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 8  
LENGTH: 61  
TYPE: PRT  
ORGANISM: Fragment 5  
US-09-828-000-8

Query Match 20.5%; Score 318; DB 4; Length 61;  
Best Local Similarity 100.0%; Pred. No. 1.7e-21;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPGTKKVHVFNYKGVKNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
DB 2 GPGTKKVHVFNYKGVKNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVESGSL 61

PRIOR APPLICATION NUMBER: 60/240,409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 947  
LENGTH: 610  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6673549 3876162CDI  
US-09-976-594-947

Query Match 37.2%; Score 577.5; DB 4; Length 610;  
Best Local Similarity 33.3%; Pred. No. 2.6e-43;  
Matches 127; Conservative 60; Mismatches 83; Indels 111; Gaps 10;

QY 6 KVHVFNYKGVKNVLINKDIRCK-----DDEFTHTLYTLIVRPDNTYEVKIDNSQVE 55  
DB 190 KLFIFRKHHPKTVGFBEKHAKPPDVLKFFTRKTHLYTLVNPDPDTTEVLVDQTVN 249

QY 56 SGLSLEDDW--DFLPPKIKDPDASKPEDWDERAKIDDPDTSKPEDW----- 99  
DB 250 KGSLEDDVVPKPKETIEDNDKKPEWDERAKIPDPSAVKPEDWDESEPAQIEDSSV 309

QY 100 -----DKPEHIPDPDAKKPEDWDEMDGEWEPVIONPEYK 135  
DB 310 KPAGWLDDEPFIPDPAEKEDDNNEDTDGEWEAPQILNPACRIGCGEWEKPPMIDNPYK 369

QY 136 GEWKPRIDNDPKYKGTWHPHIDNPEYSPDPSIYAYDNFGVLGLDLWQVSGTITFDNFI 195  
DB 370 GWRPPLVDNPNYQINSRPIKPNPDYEDDHPHILTSFSLGELNSMTSDIYFDNFII 429

QY 196 TNDAYAEFGNETW-----GVTK-----AAB----- 217  
DB 430 CSEKEVADHAAADGNRWKIMIANANKPGVLKQMAAEGHPWLWLIYLVITAGVPALITS 489

QY 218 -----KQMKDQDEQR-----LKEEEDKKKEE-----EAEKDEDED--K 254  
DB 490 FCWPRKVKKKHKTQYKTKTDCIPOTKGVLFQEEKEEKAALPKMDLEEKKQNDGEMLE 549

QY 255 DEDEDEDEKDEEDEDVPGQ 275  
DB 550 KEESEPEKSEEEIEIEGQ 570

RESULT 5  
US-08-675-816-6  
Sequence 6, Application US/08675816  
Patent No. 6171864  
GENERAL INFORMATION:  
APPLICANT: Coughlan, Sean J.  
APPLICANT: Winfrey, Jr., Ron J.  
TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND USES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry  
STREET: 701 Fifth Ave. Suite 6300  
CITY: Seattle  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/675,816  
FILING DATE: 05-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6171864tenburg, Carol

```
RESULT 7
US-09-828-000-3
; Sequence 3, Application US/09828000
; Patent No. 6596690
; GENERAL INFORMATION:
; APPLICANT: Government of the United States of America
; TITLE OF INVENTION: Vasostatin as Marrow Protectant
; FILE REFERENCE: 4239-55414
; CURRENT APPLICATION NUMBER: US/09/828,000
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Vasostatin
US-09-828-000-3
Query Match      20.5%; Score 318; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 7e-21;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKKHVIFNYGKGNVLINKDIRCKDDFTHTLYTLVRPDNTYEVKIDNSQVESGSL 60
Db 121 GPGTKKHVIFNYGKGNVLINKDIRCKDDFTHTLYTLVRPDNTYEVKIDNSQVESGSL 180

RESULT 8
US-09-828-000-4
; Sequence 4, Application US/09828000
; Patent No. 6596690
; GENERAL INFORMATION:
; APPLICANT: Government of the United States of America
; TITLE OF INVENTION: Vasostatin as Marrow Protectant
; FILE REFERENCE: 4239-55414
; CURRENT APPLICATION NUMBER: US/09/828,000
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Fragment 1
US-09-828-000-4
Query Match      15.1%; Score 235; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 4.1e-14;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPGTKKHVIFNYGKGNVLINKDIRCKDDFTHTLYTLVRPDNTYEVKIDNSQVESGSL 43
Db 19 GPGTKKHVIFNYGKGNVLINKDIRCKDDFTHTLYTLVRPDNTYEVKIDNSQVESGSL 61

RESULT 9
US-09-828-000-7
; Sequence 7, Application US/09828000
; Patent No. 6596690
; GENERAL INFORMATION:
; APPLICANT: Government of the United States of America
; TITLE OF INVENTION: Vasostatin as Marrow Protectant
; FILE REFERENCE: 4239-55414
; CURRENT APPLICATION NUMBER: US/09/828,000
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Fragment 4
US-09-828-000-7
Query Match      12.2%; Score 189; DB 4; Length 35;
```

```
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 VIFNYGKGNVLINKDIRCKDDFTHTLYTLVRPDN 43
Db 1 VIFNYGKGNVLINKDIRCKDDFTHTLYTLVRPDN 35

RESULT 10
US-08-990-114-3
; Sequence 3, Application US/08990114
; Patent No. 5932475
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,114
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0451 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 714 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 128842
US-08-990-114-3
Query Match      11.9%; Score 184.5; DB 2; Length 714;
Best Local Similarity 23.9%; Pred. No. 3.4e-08;
Matches 59; Conservative 26; Mismatches 39; Indels 123; Gaps 7;

Qy 67 PPKIKDPDASKPEDWDERAKIDFTDSKPEDWDPKPEHIPDPDAKPEDWDEMDGWEPE 126
Db 129 PAKGAKGNKAKSDSDDEDDDDSDSD-----EDEDSD-SEPEP 171
Qy 127 PVIQNPYKGEWKPERQIDNPYKGTWHPDNPYSPDSIAYDNGVLGLDLWQVKS 186
Db 172 PVV-----KG-----KQ 178
Qy 187 GTIFDNFLITNDEAYABEFGNETGWVTKAAEKQWKDKQDEEQRLEKKEEEDKKRKE 241
Db 179 GKVA-----AAAPASDEDEDEDEDEDEDEDEDEDEEAEVI 218
```







```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-
US-09-410-399-2

```

[illegible]

Search completed: October 12, 2004, 15:49:26  
Job time : 21 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 15:48:36 ; Search time 48 Seconds  
(without alignments)  
1877.162 Million cell updates/sec

Title: US-09-807-148-9

Perfect score: 1553

Sequence: 1 GPCTKKVHVFYFKGNVLI.....EEDKDEDEEDVPGQAKDEL 280

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553	100.0	417	10	US-09-906-393A-36
2	1553	100.0	417	14	Sequence 36, Appl
3	1553	100.0	417	15	Sequence 29, Appl
4	1549	99.7	416	15	Sequence 14, Appl
5	1549	99.7	416	15	Sequence 2, Appl
6	1476.5	95.1	416	14	US-10-405-588-2
7	1476.5	95.1	416	14	Sequence 2, Appl
8	1476.5	95.1	416	14	Sequence 4, Appl
9	991.5	63.8	395	15	US-10-316-253-4
10	847	54.5	421	16	US-10-316-253-6
11	844.5	54.4	442	12	US-10-369-493-6343
12	844	54.3	346	12	Sequence 6343, Ap
13	844	54.3	435	12	Sequence 46544, A
14	837.5	53.9	420	9	US-10-425-114-46869
15	835	53.8	431	16	Sequence 48930, A

Sequence 201420, A	424	16	US-10-437-963-201420
Sequence 38848, A	268	12	US-10-425-114-38848
Sequence 190956, A	391	12	US-10-424-599-190956
Sequence 190958, A	420	12	US-10-424-599-190958
Sequence 46228, A	391	12	US-10-425-114-46228
Sequence 55995, A	391	12	US-10-425-114-55995
Sequence 114914, A	442	16	US-10-437-963-114914
Sequence 201427, A	424	16	US-10-437-963-201427
Sequence 46405, A	388	12	US-10-425-114-46405
Sequence 69619, A	391	12	US-10-425-114-69619
Sequence 114860, A	667	16	US-10-437-963-114860
Sequence 277445, A	422	12	US-10-424-599-277445
Sequence 122222, A	425	16	US-10-437-963-122222
Sequence 45264, A	422	16	US-10-767-701-45264
Sequence 153808, A	418	12	US-10-424-599-153808
Sequence 70140, A	434	12	US-10-425-114-70140
Sequence 65495, A	439	12	US-10-425-114-65495
Sequence 57598, A	434	12	US-10-425-114-57598
Sequence 44621, A	372	12	US-10-425-114-44621
Sequence 54997, A	372	12	US-10-425-114-54997
Sequence 129, App	591	14	US-10-205-194-129
Sequence 1095, Ap	592	16	US-10-408-765A-1095
Sequence 56908, A	207	12	US-10-425-114-56908
Sequence 61049, A	610	14	US-10-205-823-68
Sequence 46, Appl	184	16	US-10-767-701-61049
Sequence 55857, A	390	15	US-10-161-927-46
Sequence 4044, A	167	16	US-10-767-701-55857
Sequence 47162, A	536	12	US-10-425-114-40444
Sequence 146333, A	533	16	US-10-767-701-47162
	537	16	US-10-437-963-146333

ALIGNMENTS

RESULT 1  
US-09-906-393A-36  
; Sequence 36, Application US/09906393A  
; Publication No. US20030039970A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Zhou  
; APPLICANT: Xiao, Wuhan  
; TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED  
; FILE REFERENCE: 1720-1-001CIP  
; CURRENT APPLICATION NUMBER: US/09/906,393A  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/218,761  
; PRIOR FILING DATE: 2000-07-17  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36  
; LENGTH: 417  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-906-393A-36

Query Match	100.0%;	Score 1553;	DB 10;	Length 417;
Best Local Similarity	100.0%;	Pred. No. 1e-106;		
Matches 280;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GPCTKKVHVFYFKGNVLI	IRCKDDEFFHLYTLIVRPDNTYEVKIDNSQVSGSLE	60
Db	138	GPCTKKVHVFYFKGNVLI	IRCKDDEFFHLYTLIVRPDNTYEVKIDNSQVSGSLE	197
Qy	61	DDWDFLPPKKIKDPDASKPEDMDERAKIDDPDTSKPEDMDKPEHIPDDPAKPEDWDEEM	120	
Db	198	DDWDFLPPKKIKDPDASKPEDMDERAKIDDPDTSKPEDMDKPEHIPDDPAKPEDWDEEM	257	
Qy	121	DGWESEPVTONPEYKGEWKPRQIDNPDYKGTWHPIDNPEYSPDPSIYAYDNFVGLGD	180	
Db	258	DGWESEPVTONPEYKGEWKPRQIDNPDYKGTWHPIDNPEYSPDPSIYAYDNFVGLGD	317	
Qy	181	LWQVSGTTFPNFLITNDEAYAEFGNETGWVTGVTAAEKQMDKQDEQRKLBSEEDKKRK	240	



```

RESULT 5
US-10-405-588-2
; Sequence 2, Application US/10405588
; Publication No. US20030216299A1
; GENERAL INFORMATION:
; APPLICANT: Government of the United States of America
; TITLE OF INVENTION: Vasostatin as Marrow Protectant
; FILE REFERENCE: 4239-55414
; CURRENT APPLICATION NUMBER: US/10/405,588
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US/09/828,000
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Calreticulin
US-10-405-588-2

Query Match      99.7%; Score 1549; DB 15; Length 416;
Best Local Similarity 100.0%; Pred. No. 2e-106;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  GPGTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 60
Db      138  GPGTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 197

Qy      61  DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDPKEHIPDPDAKPPEDWDEEM 120
Db      198  DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDPKEHIPDPDAKPPEDWDEEM 257

Qy      121  DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTIAYDNFVGLGLD 180
Db      258  DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTIAYDNFVGLGLD 317

Qy      181  LQVKSGTTFDNFLTNDAYAEFFGNETGWTKAAEKQMKDKQDEEQRLEKEEEDKKRK 240
Db      318  LQVKSGTTFDNFLTNDAYAEFFGNETGWTKAAEKQMKDKQDEEQRLEKEEEDKKRK 377

Qy      241  EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
Db      378  EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 416

RESULT 6
US-10-316-253-2
; Sequence 2, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-2

Query Match      95.1%; Score 1476.5; DB 14; Length 416;
Best Local Similarity 94.6%; Pred. No. 4.6e-101;
Matches 265; Conservative 10; Mismatches 4; Indels 1; Gaps 1;

Qy      1  GPGTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 60
Db      138  GPGTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 197

Qy      61  DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDPKEHIPDPDAKPPEDWDEEM 120
Db      198  DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDPKEHIPDPDAKPPEDWDEEM 257

Qy      121  DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTIAYDNFVGLGLD 180
Db      258  DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTIAYDNFVGLGLD 317

Qy      181  LQVKSGTTFDNFLTNDAYAEFFGNETGWTKAAEKQMKDKQDEEQRLEKEEEDKKRK 240
Db      318  LQVKSGTTFDNFLTNDAYAEFFGNETGWTKAAEKQMKDKQDEEQRLEKEEEDKKRK 377

Qy      241  EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
Db      378  EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 416

RESULT 7
US-10-316-253-4
; Sequence 4, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-4

Query Match      95.1%; Score 1476.5; DB 14; Length 416;
Best Local Similarity 94.6%; Pred. No. 4.6e-101;
Matches 265; Conservative 10; Mismatches 4; Indels 1; Gaps 1;

Qy      1  GPGTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 60
Db      138  GPGTKVHVIFNYKGNVLINKDIRCKDDEFTHTLYLIVRPDNTYEVKIDNSQVESGSL 197

Qy      61  DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDPKEHIPDPDAKPPEDWDEEM 120
Db      198  DDWDFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDPKEHIPDPDAKPPEDWDEEM 257

Qy      121  DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTIAYDNFVGLGLD 180
Db      258  DGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWHPIDNPYSPDPSTIAYDNFVGLGLD 317

Qy      181  LQVKSGTTFDNFLTNDAYAEFFGNETGWTKAAEKQMKDKQDEEQRLEKEEEDKKRK 240
Db      318  LQVKSGTTFDNFLTNDAYAEFFGNETGWTKAAEKQMKDKQDEEQRLEKEEEDKKRK 377

Qy      241  EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
Db      378  EEEEAEDKEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 416

RESULT 8
US-10-316-253-6
; Sequence 6, Application US/10316253
```



GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 46869  
; LENGTH: 442  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73073A05\_FLI.pep  
US-10-425-114-46869

Query Match 54.4%; Score 844.5; DB 12; Length 442;  
Best Local Similarity 53.0%; Pred. No. 2.3e-54;  
Matches 149; Conservative 49; Mismatches 78; Indels 5; Gaps 2;

Qy 1 GPGTKKHVIFNYKGNVLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSGSLE 60  
Db 166 GYSTKKVHTILTCKGNHLIKKQVPCETDQLTHVYTLIRPDATYSILIDNEKQTSIY 225

Qy 61 DWDWFLPPKTKDPPDASKPEDWDERAKIDDDPTDSKPEDWDK-PHIPPDPDAKPEDWDEE 119  
Db 226 EHWDLILPPKTKDPPDASKPEDWDDKEYIPDPEDKKPEGYDDIPKEIPDPDAKPEDWDE 285

Qy 120 MDGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWTHPEIDNPYSPDPSIYAYDNFVGL 179  
Db 286 EDGEWTAFTIPNPYKGPWKQKIKPNYQGWKAPMIDNPDYKDDPYIYAFDSLKYIGI 345

Qy 180 DLQVKSGLTIFDNLITNDEAYAEFGNETGWVTKAAEQKMDKQDEEQLKEEEDKKR 239  
Db 346 ELQVKSGLTIFDNLITNDEAYAEFGNETGWVTKAAEQKMDKQDEEQLKEEEDKKR 401

Qy 240 KEEEAEDKEDDE 280  
Db 402 GGDEDDDDLEDE 442

RESULT 12  
US-10-425-114-48930  
; Sequence 48930, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 48930  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3150-081-A9\_FLI.pep  
US-10-425-114-48930

Query Match 54.3%; Score 844; DB 12; Length 346;  
Best Local Similarity 53.7%; Pred. No. 1.9e-54;  
Matches 149; Conservative 49; Mismatches 78; Indels 5; Gaps 2;

Matches 154; Conservative 46; Mismatches 71; Indels 16; Gaps 5;

Qy 1 GPGTKKHVIFNYKGNVLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSGSLE 60  
Db 69 GYSTKKVHTILTCKGNHLIKKQVPCETDQLTHVYTLIRPDATYSILIDNEKQTSIY 128

Qy 61 DWDWFLPPKTKDPPDASKPEDWDERAKIDDDPTDSKPEDWDK-PHIPPDPDAKPEDWDEE 119  
Db 129 EHWDLILPPKTKDPPDASKPEDWDDKEYIPDPEDKKPEGYDDIPKEIPDPDAKPEDWDE 188

Qy 120 MDGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWTHPEIDNPYSPDPSIYAYDNFVGL 179  
Db 189 EDGEWTAFTIPNPYKGPWKQKIKPNYQGWKAPMIDNPDYKDDPYIYAFDSLKYIGI 248

Qy 180 DLQVKSGLTIFDNLITNDEAYAEFGNETGWVTKAAEQKMDKQDEEQLKEEEDKKR 239  
Db 249 ELQVKSGLTIFDNLITNDEAYAEFGNETGWVTKAAEQKMDKQDEEQLKEEEDKKR 304

Qy 240 KEEEAEDKEDDE 280  
Db 305 GGDEDDDDLEDE 346

RESULT 13  
US-10-425-114-70304  
; Sequence 70304, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 70304  
; LENGTH: 435  
; TYPE: PRT  
; ORGANISM: Zea mays subsp. mexicana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE102C08\_FLI.pep  
US-10-425-114-70304

Query Match 54.3%; Score 844; DB 12; Length 435;  
Best Local Similarity 53.7%; Pred. No. 2.4e-54;  
Matches 154; Conservative 46; Mismatches 71; Indels 16; Gaps 5;

Qy 1 GPGTKKHVIFNYKGNVLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSGSLE 60  
Db 158 GYSTKKVHTILTCKGNHLIKKQVPCETDQLTHVYTLIRPDATYSILIDNEKQTSIY 217

Qy 61 DWDWFLPPKTKDPPDASKPEDWDERAKIDDDPTDSKPEDWDK-PHIPPDPDAKPEDWDEE 119  
Db 218 EHWDLILPPKTKDPPDASKPEDWDDKEYIPDPEDKKPEGYDDIPKEIPDPDAKPEDWDE 277

Qy 120 MDGEWEPPIQNPYKGEWKPRQIDNPDKYKGTWTHPEIDNPYSPDPSIYAYDNFVGL 179  
Db 278 EDGEWTAFTIPNPYKGPWKQKIKPNYQGWKAPMIDNPDYKDDPYIYAFDSLKYIGI 337

Qy 180 DLQVKSGLTIFDNLITNDEAYAEFGNETGWVTKAAEQKMDKQDEEQLKEEEDKKR 239  
Db 338 ELQVKSGLTIFDNLITNDEAYAEFGNETGWVTKAAEQKMDKQDEEQLKEEEDKKR 393

Qy 240 KEEEAEDKEDDE 280  
Db 394 GGDEDDDDLEDE 435

## RESULT 14

US-09-844-006A-2  
 ; Sequence 2, Application US/09844006A  
 ; Patent No. US20020083496A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wyatt, Sarah  
 ; APPLICANT: Tseu, Pei-Lan  
 ; APPLICANT: Robertson, Dominique  
 ; APPLICANT: Boss, Wendy  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH INCREASES IN CALCIUM STORES  
 ; FILE REFERENCE: 5051.503  
 ; CURRENT APPLICATION NUMBER: US/09/844.006A  
 ; CURRENT FILING DATE: 2001-04-30  
 ; PRIOR APPLICATION NUMBER: 60/200,233  
 ; PRIOR FILING DATE: 2000-04-28  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 420  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 US-09-844-006A-2

Query Match 53.9%; Score 837.5; DB 9; Length 420;

Best Local Similarity 52.9%; Pred. No. 7.1e-54;

Matches 147; Conservative 49; Mismatches 77; Indels 5; Gaps 2;

```

QY 4 TKKVHVIFNYKGNVLINKDIRCKDDSEFTHLYTLIVRPDNTYEVKIDNSQVESGSLEDDW 63
DB 147 TKVHTLTGDKGNHLIKGVPCQTDQTHVYTLIRPDATYSLIDNEEKTCGSIYEHW 206

QY 64 DFLPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWK- PEHIPDPDAKPEDWDEMDG 122
DB 207 DILPPKKIKDPEAKPEDWDKEYIPDPEDKKPGYDDIPKEIPDPDAKPEDWDEEDG 266

QY 123 EWEPPVTONPEYKGEWKPQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFVGLDLW 182
DB 267 EWTAPTIPNPEYKGPWKQKIKPNYQGWKAPMIDNPDFKDDPYIYAFDSLKYIGIELW 326

QY 183 QVKSGLTFDNLITNDAYAEFGNETGWVTKAAEQKQDKQDEEQRLEKEEEDKKRKEE 242
DB 327 QVKSGLTFDNLITDDPALAKTFAEETWKGKKEAKAFADEAK- ----KKEEEDAAKGGD 382

QY 243 EAEDKEDDEKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
DB 383 DEDDDLEDEDEDEKADSDAEDGKDSDEKXDEL 420

```

## RESULT 15

US-10-767-701-44015  
 ; Sequence 44015, Application US/10767701  
 ; Publication No. US20040172684A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
 ; FILE REFERENCE: 38-21(53535)B  
 ; CURRENT APPLICATION NUMBER: US/10/767,701  
 ; CURRENT FILING DATE: 2004-01-29  
 ; NUMBER OF SEQ ID NOS: 63128  
 ; SEQ ID NO 44015  
 ; LENGTH: 431  
 ; TYPE: PRT  
 ; ORGANISM: Sorghum bicolor  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(431)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C10317\_1.pbp  
 US-10-767-701-44015

Query Match 53.8%; Score 835; DB 16; Length 431;  
 Best Local Similarity 54.3%; Pred. No. 1.1e-53;  
 Matches 159; Conservative 35; Mismatches 75; Indels 24; Gaps 5;

```

QY 1 GPGTKKVHVIFNYKGNVLINKDIRCKDDSEFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 60
DB 150 GYATKKVHAILTNGKNHLIKKELPGETDQTHVYTLIRPDATYSLIDNEKQSGSIY 209

QY 61 DWDDELPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWK- PEHIPDPDAKPEDWDEE 119
DB 210 DWDDELPPKKIKDPDASKPEDWDDKEYIPDPEDKKPGYDDIPKEIPDPDAKPEDWDE 269

QY 120 MDGEWEPVTONPEYKGEWKPQIDNPDKYKGTWHPIDNPESPDPSIYAYDNFVGLGL 179
DB 270 EDGEWTAPTIPNPEYKGPWKQKIKPNYQGWKAPLIDNPDFKDDPYIYAFDSLKHGI 329

QY 180 DLWQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEQKQDKQDEEQRLEKEEEDKKR 239
DB 330 ELWQVKSGLTFDNLITDDPEYAKKLAVETWKGKDAEKAADF-----EAEKKR 378

QY 240 KEE-----EEAEDKEDDE--DKDEDEDEDE-----KEEEDDEEVPGQAKDEL 280
DB 379 LEECFSLXESASSKODDDDLVDDEDEDDADDKADHSDTDEAEDSEEAKHDEL 431

```

Search completed: October 12, 2004, 15:54:17

Job time : 50 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 15:46:05 ; Search time 17 Seconds  
(without alignments)  
1584.331 Million cell updates/sec

Title: US-09-807-148-9  
Perfect score: 1553  
Sequence: 1 GFGTKKHVIFNYKGNVLI.....BEDKEDEEDVPGQAKDEL 280

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Query %	Length	ID	Description
1	1553	100.0	417	1	A37047	calreticulin precu
2	1487.5	95.8	416	1	S08763	calreticulin precu
3	1476.5	95.1	416	2	JH0819	calreticulin precu
4	1457.5	93.9	418	1	A34154	calreticulin precu
5	1442	92.9	400	2	S43376	calreticulin, brai
6	1432	92.2	421	2	S36799	calreticulin precu
7	1280.5	82.5	411	2	S29129	calreticulin precu
8	1249.5	80.5	384	2	S29130	calreticulin (clon
9	1241.5	79.9	419	2	JH1243	calreticulin precu
10	1184	75.0	405	1	JH0795	calreticulin precu
11	1031.5	66.4	406	2	A56637	calreticulin homol
12	991.5	63.8	395	2	S25951	calreticulin precu
13	944.5	60.8	336	2	A32507	41K larval antigen
14	856.5	55.2	415	2	T10172	calreticulin - cas
15	844	54.3	421	2	S58170	calreticulin precu
16	818.5	52.7	416	2	T15968	calreticulin call
17	817.5	52.6	389	2	T03691	calreticulin - com
18	817.5	52.6	393	1	A48573	calreticulin autoa
19	814.5	52.4	416	2	T14554	calreticulin - bee
20	813	52.4	422	2	T07841	probable calreticu
21	812	52.3	415	2	T05703	calreticulin - bar
22	812	52.3	412	2	T05705	calreticulin - bar
23	810.5	52.2	425	2	C96605	calreticulin (ctrl
24	801.5	51.6	444	2	H85224	hypothetical prote
25	643	41.4	622	2	S71342	calnexin precursor
26	633	40.8	591	2	B54354	calnexin precursor
27	639	40.5	591	2	C54354	calnexin precursor
28	628	40.4	592	2	I53260	calnexin - human
29	628	40.4	592	2	A46673	calnexin precursor

30	624	40.2	593	1	A37273	calnexin precursor
31	587	37.8	356	2	A46164	calnexin - human
32	555	35.7	611	2	A54086	calnexin-t - mouse
33	554	35.7	611	2	A53418	calmeglin precursor
34	539.5	34.7	582	2	A46637	calnexin homolog S
35	531	34.2	546	2	T06415	calnexin - soybean
36	501.5	32.3	428	2	T03251	calnexin - maize
37	501.5	32.3	530	2	JN0597	calnexin-like prot
38	477.5	30.7	532	2	T49873	calnexin homolog -
39	472.5	30.4	560	2	S56142	calcium-binding pr
40	472.5	30.4	619	2	S40938	hypothetical prote
41	458.5	29.5	297	2	S70552	calnexin homolog C
42	451.5	29.1	540	2	T10892	probable calnexin
43	299.5	19.3	502	2	S29347	calnexin homolog Y
44	208.5	13.4	792	2	T42963	hypothetical prote
45	193	12.4	798	2	T33022	hypothetical prote

## ALIGNMENTS

## RESULT 1

A37047  
calreticulin precursor - human  
N;Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding prote:  
C;Species: Homo sapiens (man)  
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 18-Feb-2000  
C;Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075  
R;McCaulliffe, D.P.; Yang, Y.S.; Wilson, J.; Sontheimer, R.D.; Capra, J.D.  
J. Biol. Chem. 267, 2557-2562, 1992  
A;Title: The 5'-flanking region of the human calreticulin gene shares homology with the  
A;Reference number: A42330; MUID:92129342; PMID:1733953  
A;Accession: A42330  
A;Molecule type: DNA  
A;Residues: 1-417 <MC2>  
A;Note: sequence extracted from NCBI backbone (NCBIN:78537, NCBI:78536)  
R;McCaulliffe, D.P.; Lux, F.A.; Liew, T.S.; Sanz, I.; Hanke, J.; Newkirk, M.M.; Bachinski  
J. Clin. Invest. 85, 1379-1391, 1990  
A;Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/SS-  
A;Reference number: A37047; MUID:90237213; PMID:2332496  
A;Accession: A37047  
A;Molecule type: mRNA  
A;Residues: 1-417 <MC2>  
A;Cross-references: GB:M32294; NID:g337486; PIDN:AAA36582.1; PID:g337487  
A;Note: the authors translated the codon GTA for residue 349 as Tyr  
R;Rokach, J.A.; Haseby, J.A.; Meilof, J.F.; Smeenk, R.J.; Umansh, T.R.; Greene, B.M.;  
J. Immunol. 147, 3031-3039, 1991  
A;Title: Characterization of the autoantigen calreticulin.  
A;Reference number: A46452; MUID:92013129; PMID:1919005  
A;Accession: A46452  
A;Molecule type: mRNA  
A;Residues: 1-417 <R0K>  
A;Cross-references: GB:M84739; NID:g179881; PIDN:AAA51916.1; PID:g179882  
A;Note: sequence extracted from NCBI backbone (NCBIN:60749, NCBI:60750)  
R;Liew, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.  
J. Clin. Invest. 82, 96-101, 1988  
A;Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence of  
A;Reference number: A28812; MUID:88273610; PMID:3260607  
A;Accession: A28812  
A;Molecule type: protein  
A;Residues: 18-41 <LIE>  
A;Note: 18-Ala was also found  
R;Dupuis, M.; Schaefer, E.; Krause, K.H.; Tschopp, J.  
J. Exp. Med. 177, 1-7, 1993  
A;Title: The calcium-binding protein calreticulin is a major constituent of lytic granul  
A;Reference number: PH1525; MUID:93115648; PMID:8418194  
A;Accession: PH1525  
A;Molecule type: protein  
A;Residues: 18-27 <DUF>  
A;Experimental source: LAK cell  
R;Rojiani, M.V.; Finlay, S.B.; Gray, V.; Dedhar, S.  
Biochemistry 30, 9859-9865, 1991  
A;Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (calr



RESULT 5  
S43376  
calretic  
C;Specie  
C;Date:

```

F:193/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match          92.2%; Score 1432; DB 2; Length 421;
Best Local Similarity 92.2%; Pred. No. 5.7e-80;
Matches 259; Conservative 12; Mismatches 8; Indels 2; Gaps 2;

QY      1  GPGTKKHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEVKIDNSQVESGSLE 60
Db      142 GPGTKKHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEVKIDNSQVESGSLE 201
QY      61  DDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDKPEHI PDPDAKKPEDWDEEM 120
Db      202 DDWDFLPPKKIKDPDAAKPEDWDRAKIDDPDTSKPEDWDKPEHI PDPDAKKPEDWDEEM 261
QY      121 DGWESEPVIONPEYKGEWKPRQIDNPDKYGTWTHPEIDNPESYSPDPSIYAYDNFGVLGLD 180
Db      262 DGWESEPLIONPEYKGEWKPRQIDNPDKYGTWTHPEIDNPESYSPDPSIYAYDNFGVLGLD 321
QY      181 LWQVKSGTIFDNFLITNDAYAEAFGNETWGTVKAEEKQMKQDEQRLEKEEEDKKRK 240
Db      322 LWQVKSGTIFDNFLITNDAYAEAFGNETWGTVKAEEKQMKQDEQRLEHEEEKKRK 381
QY      241 EEEAEADKDDDKDEDEDEDEDEDEED-VPGQAKDEL 280
Db      382 EEEAE- KDDDEKDEDEDEDEDEDEDEAAAAQAKDEL 421

RESULT 7
S29129
calreticulin precursor (clone 3) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: S29129
R:Reves, S.; Zorzato, F.; Pozzan, T.
Biochem. J. 287, 579-581, 1992
A:Title: Identification of calreticulin isoforms in the central nervous system.
A:Reference number: S29129; MUID:93074997; PMID:1445218
A:Accession: S29129
A:Molecule type: mRNA
A:Residues: 1-411 <TRE>
A:Cross-references: EMBL:X67597; NID:G64608; PIDN:CAA47866.1; PID:G64609
C:Superfamily: calreticulin
C:Keywords: glycoprotein
F:1-12/Domain: signal sequence (fragment) #status predicted <SIG>
F:13-411/Product: calreticulin #status predicted <MAT>
F:408-411/Region: endoplasmic reticulum retention signal
F:339/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          82.5%; Score 1280.5; DB 2; Length 411;
Best Local Similarity 80.1%; Pred. No. 8.6e-71;
Matches 225; Conservative 31; Mismatches 22; Indels 3; Gaps 2;

QY      1  GPGTKKHVIFNYKGNVLINKDIRCKDDETHLYTLIVRPNTYEVKIDNSQVESGSLE 60
Db      133 GPPTKKHVIFQYKXKNLQINKDIRCKDDSFTHLYTLIVRPNTYEVKIDNSQVESGSLE 192
QY      61  DDWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDKPEHI PDPDAKKPEDWDEEM 120
Db      193 DDWDFLPPKKIKDPEAKPDWDERPKIDDPEDKKPEDWKEPHEI PDPDAKPEDWDEEM 252
QY      121 DGWESEPVIONPEYKGEWKPRQIDNPDKYGTWTHPEIDNPESYSPDPSIYAYDNFGVLGLD 180
Db      253 DGWESEPVIONPEYKGEWKPRQIDNPDKYGTWTHPEIDNPESYSPDPSIYAYDNFGVLGLD 312
QY      181 LWQVKSGTIFDNFLITNDAYAEAFGNETWGTVKAEEKQMKQDEQRLEKEEEDKKRK 240
Db      313 LWQVKSGTIFDNFLITNDKHAAYEYGNETWGTVKAEEKQMKQDEDRKKQEESEKTRK 372
QY      241 EEE-EADKEDDKDEDEDEDEDEDEDEEDVPGQAKDEL 280
Db      373 EEEFQEEDEDDDEKEEKEEKEEKEEEDSEETP-LKDEL 411

```

RESULT 8  
 S29130  
 C:Species: Xenopus laevis (African clawed frog (fragment))  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 13-Aug-1999  
 C:Accession: S29130; T01068  
 R:Preves, S.; Zorzato, F.; Pozzan, T.  
 Biochem. J. 287, 579-581, 1992  
 A:Title: Identification of calreticulin isoforms in the central nervous system.  
 A:Reference number: S29129; MUID:93074997; PMID:1445218  
 A:Accession: S29130  
 A:Molecule type: mRNA  
 A:Residues: 1-384 <TR>  
 A:Cross-references: EMBL:X67598  
 A:Accession: T01068  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-339, <XTGR> <TR>  
 A:Cross-references: EMBL:X67598; NID:G64610; PIDN:CAA47867.1; PID:G64611  
 A:Experimental source: CNS  
 C:Superfamily: calreticulin  
 C:Keywords: glycoprotein  
 F:381-384/Region: endoplasmic reticulum retention signal  
 F:316/Binding site: carbohydrate (Asn) (covalent) #status predicted  
  
 Query Match 80.5%; Score 1249.5; DB 2; Length 384;  
 Best Local Similarity 78.2%; Pred. No. 6.1e-69;  
 Matches 219; Conservative 30; Mismatches 26; Indels 5; Gaps 1;  
  
 QY 1 GPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
 DB 110 GPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 169  
  
 QY 61 DWDDELPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKKPEDWDEEM 120  
 DB 170 EDWDFLPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKKPEDWDEEM 229  
  
 QY 121 DGEWPPVITNPEYKGNPKQIDNPDKYKGTWHPEDINPEYSDPSIAYDNGVLGLD 180  
 DB 230 DGEWPPVITNPEYKGNPKQIDNPDKYKGTWHPEDINPEYSDPSIAYDNGVLGLD 289  
  
 QY 181 LMQVSGTIFDNLITNDEAYAEFGNETGWTKAAEKQKDKODEEORLKEEEDKKRK 240  
 DB 290 LMQVSGTIFDNLITNDEAYAEFGNETGWTKAAEKQKDKODEEORLKEEEDKKRK 349  
  
 QY 241 EEEAEEDKDDDE 280  
 DB 350 EEEAEEDKDDDE 384  
  
 RESULT 9  
 S71343  
 C:Species: Rana rugosa (Korean frog)  
 C:Date: 29-Jan-1998 #sequence\_revision 13-Feb-1998 #text\_change 20-Jun-2000  
 C:Accession: S71343  
 R:Yamamoto, S.; Nakamura, M.  
 FEBS Lett. 387, 27-32, 1996  
 A:Title: Calnexin: its molecular cloning and expression in the liver of the frog, Rana  
 A:Reference number: S71342; MUID:96234004; PMID:8654561  
 A:Accession: S71343  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-419 <YAM>  
 A:Cross-references: EMBL:D78589; NID:g1514956; PIDN:BAAL1425.1; PID:g1514957  
 C:Superfamily: calreticulin  
 C:Keywords: calcium binding; endoplasmic reticulum  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-419/Product: calreticulin #status predicted <MAT>  
 F:205-213/Region: nuclear location signal  
 F:415-418/Region: endoplasmic reticulum retention signal  
  
 Query Match 79.9%; Score 1241.5; DB 2; Length 419;

Best Local Similarity 76.9%; Pred. No. 2e-68;  
 Matches 216; Conservative 34; Mismatches 30; Indels 1; Gaps 1;  
  
 QY 1 GPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
 DB 139 GPGTKKVVIFNYKGNVLINKDIRKADVVSHTLYTLIVRPDNTYEVKIDNSQVESGSL 198  
  
 QY 61 DWDDELPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKKPEDWDEEM 120  
 DB 199 DWDDELPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKKPEDWDEEM 258  
  
 QY 121 DGEWPPVITNPEYKGNPKQIDNPDKYKGTWHPEDINPEYSDPSIAYDNGVLGLD 180  
 DB 259 DGEWPPVITNPEYKGNPKQIDNPDKYKGTWHPEDINPEYSDPSIAYDNGVLGLD 318  
  
 QY 181 LMQVSGTIFDNLITNDEAYAEFGNETGWTKAAEKQKDKODEEORLKEEEDKKRK 240  
 DB 319 LMQVSGTIFDNLITNDEAYAEFGNETGWTKAAEKQKDKODEEORLKEEEDKKRK 378  
  
 QY 241 EEEAEEDKDDDE 280  
 DB 379 EEEAEEDKDDDE 419  
  
 RESULT 10  
 JH0795  
 C:Species: Aplysia California (California sea hare)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: JH0795; B31409; F60977  
 R:Kennedy, T.E.; Kuhl, D.; Barzilai, A.; Sweatt, J.D.; Kandel, E.R.  
 Neuron 9, 1013-1024, 1992  
 A:Title: Long-term sensitization training in aplysia leads to an increase in calreticulin.  
 A:Reference number: JH0795; MUID:93098937; PMID:1463604  
 A:Accession: JH0795  
 A:Molecule type: mRNA  
 A:Residues: 1-405 <KEN>  
 A:Cross-references: GS:S51239; NID:G262053; PIDN:AAB24569.1; PID:G262054  
 A:Experimental source: abdominal ganglion and antral nervous system  
 R:Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.; Sweatt, J.D.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 7008-7012, 1988  
 A:Title: Sequencing of proteins from two-dimensional gels by using in situ digestion and  
 A:Reference number: A94207; MUID:88320566; PMID:3413132  
 A:Accession: B31409  
 A:Molecule type: protein  
 A:Residues: 'X', 17-28, 'X', 30-31 <K2>  
 R:Sweatt, J.D.; Kennedy, T.E.; Wager-Smith, K.; Gawinowicz, M.A.; Barzilai, A.; Karl, K.  
 Electrophoresis 10, 152-157, 1989  
 A:Title: Development of a database of amino acid sequences for proteins identified and i  
 A:Reference number: A60977; MUID:89276264; PMID:2731514  
 A:Accession: F60977  
 A:Molecule type: protein  
 A:Residues: 'X', 17-28, 'X', 30-31 <SWE>  
 C:Superfamily: calreticulin  
 C:Keywords: calcium binding; endoplasmic reticulum  
 F:1-15/Domain: signal sequence #status predicted <SIG>  
 F:16-405/Product: calreticulin #status experimental <MAT>  
 F:402-405/Region: endoplasmic reticulum retention signal  
  
 Query Match 75.0%; Score 1164; DB 1; Length 405;  
 Best Local Similarity 72.9%; Pred. No. 9.9e-64;  
 Matches 204; Conservative 29; Mismatches 39; Indels 8; Gaps 2;

QY 1 GPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 60  
 DB 134 GPGTKKVVIFNYKGNVLINKDIRCKDDETHLYTLIVRPDNTYEVKIDNSQVESGSL 193  
  
 QY 61 DWDDELPPKKIKDPDASKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKKPEDWDEEM 120  
 DB 194 ADWDFLPARTIPDPDAKKPEDWDERAKIDDDPTDSKPEDWDXPEHIPDPDAKKPEDWDEEM 253

S25851  
calreticulin precursor - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 06-Jan-1994 #sequence\_revision 10-Nov-1995 #text\_change 05-Nov-1999  
C:Accession: S25851; T33996  
R:Smith, M.J.  
DNA Seq. 2, 235-240, 1992  
A>Title: A C. elegans gene encodes a protein homologous to mammalian calreticulin.  
A:Reference number: S25851; MUID:92325978; PMID:1627827  
A:Accession: S25851  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-395 <SMI>  
A:Cross-references: EMBL:X59589; NID:g6693; PIDN:CAA42159.1; PID:g6694  
R:Bauer, C.; Courtney, L.; Laplant, Y.  
submitted to the EMBL Data Library, February 1999  
A:Description: The sequence of C. elegans cosmid Y38A10A.  
A:Reference number: Z21453  
A:Accession: T33996  
A>Status: Preliminary; translated from GE/EMBL/DDBB  
A:Molecule type: DNA  
A:Residues: 1-395 <BAU>  
A:Cross-references: EMBL:AF125963; PIDN:AAD14746.1; GSPDB:GN00023; CESP:Y38A10A.5  
A:Experimental source: strain Bristol N2; clone Y38A10A  
C:Genetics:  
A:Gene: CESP:Y38A10A.5  
A:Map position: 5  
A:Insertions: 107/3; 315/3  
C:Superfamily: calreticulin  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:392-395/Region: endoplasmic reticulum retention signal

Query Match            63.8%; Score 991.5; DB 2; Length 395;  
Best Local Similarity 66.0%; Pred. No. 2.8e-53;  
Matches 179; Conservative 35; Mismatches 53; Indels 1; Gaps 1;

Qy 1 GPOTKKVHVIFNYKGRNVLINKDIRCKDDEFTHLTYLTIVRPDNTYEVKIDNSQVESGSLE 60  
     |||.....  
Db 134 GP-TRRVHVLNYKGKNLIKKEITCKSDLTHTLYTLILNSDNTYEVKIDGESAQTGSLE 192  
     |||.....

Qy 61 DWDWFLPPKKIKDPDASKEDWDERAKIDPTDSKEDWDKPHEIPDPAKPEDWDEEM 120  
     |||.....  
Db 193 EDMDLLPARKIKDPDAKPEDWDEREYIDAEAPKEDWEKPEHIPDPAKPEDWDDDEM 252  
     |||.....

Qy 121 DGWEPPVTIQNPYKGEWKPRQINDPYKGTWHPEIDNPESPDPSIYAYDNFGVLGLD 180  
     |||.....  
Db 253 DGWEPPMIDNPYKGEWKPKQLKNPAYKGVKHHPLENPEYTPDDELVSYSWGAIQFD 312  
     |||.....

Qy 181 LWQVKSQTTFDLINDRAYAAEFENETVGTVTKAAEKQMKOKQDEBQRLKEEBEDKKRK 240  
     |||.....  
Db 313 LWQVKSQTTFDMIIITDSVEEAHAHAETFDKLTKVEKEKKADETRKABEARAKAE 372  
     |||.....

Qy 241 EEEEAEDKDEDKDEDEDEDE 262  
     |||.....  
Db 373 EEKEAKDDDEEEKDEEGHDE 394  
     |||.....

RESULT 13  
A32507  
41k larval antigen - nematode (Onchocerca volvulus) (fragment)  
C:Species: Onchocerca volvulus  
C>Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 12-Apr-1995  
R:Accession: A32507; A28813  
R:Unnasch, T.R.; Gallin, M.Y.  
J. Clin. Invest. 82, 262-269, 1988  
A>Title: Isolation and characterization of expression cDNA clones encoding antigens  
A:Reference number: A92769; MUID:88273584; PMID:2455736  
A:Accession: A32507  
A:Molecule type: mRNA  
A:Residues: 1-336 <UNN>  
C:Superfamily: calreticulin

Query Match            60.8%; Score 944.5; DB 2; Length 336;

Best Local Similarity 65.4%; Pred. No. 1.7e-50;  
Matches 166; Conservative 32; Mismatches 49; Indels 7; Gaps 2;

QY 1 GPCTKKVHVFYKGNVILNKIDIRCKDDEFHLYTLVIRPDNTYEVKIDNSQVESGSL 60  
Db 84 GPCTKKVHVFYKGNVILNKIDIRCKDDEFHLYTLVIRPDNTYEVKIDNSQVESG 143

QY 61 DWDVFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDW 120  
Db 144 ADWDFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDW 203

QY 121 DGEWEPVTONPEYKGEWKPQIDNPDPYKGTWVHPEIDNPSPDPSIYVNDNGVLGLD 180  
Db 204 DGEWEPVNDPEYKGEWKPQKQNPAYKGIWVHPEIPIYTPDNDLYVYDDIGAGFD 263

QY 181 LWQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKMDKQDEORLKEEEDK 236  
Db 264 LWQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKMDKQDEORLKEEEDK 322

QY 237 --KKRKEEEDK 248  
Db 323 KMKKKRANRKKKK 336

## RESULT 14

T10172  
calreticulin - castor bean  
C;Species: Ricinus communis (castor bean)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C;Accession: T10172  
R;Coughlan, S.J.; Hastings, C.; Winfrey Jr., R.  
Plant Mol. Biol. 34, 897-911, 1997  
A;Title: Cloning and characterization of the calreticulin gene from Ricinus communis L.  
A;Reference number: Z16973; MUID:97435975; PMID:9290642  
A;Accession: T10172  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-415 <COU>  
A;Cross-references: EMBL:U74630; NID:g1658196; PIDN:AAB71419.1; PID:g1658197  
C;Superfamily: calreticulin  
C;Keywords: calcium binding; endoplasmic reticulum  
F;412-415/Region: endoplasmic reticulum retention signal

Query Match 55.2%; Score 856.5; DB 2; Length 415;  
Best Local Similarity 55.3%; Pred. No. 4.6e-45;  
Matches 157; Conservative 36; Mismatches 80; Indels 11; Gaps 4;

QY 1 GPCTKKVHVFYKGNVILNKIDIRCKDDEFHLYTLVIRPDNTYEVKIDNSQVESGSL 60  
Db 139 GYSTKKVHAILNVDNTHLKEVFCETDQLTHVTLVIRPDATYSLIDNVEKQTGS 198

QY 61 DWDVFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDW 119  
Db 199 TDWDLFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDW 258

QY 120 MDGEWEPVTONPEYKGEWKPQIDNPDPYKGTWVHPEIDNPSPDPSIYVNDNGVLGL 179  
Db 259 EDGEWTAFTIAPNPYKGPWKPKIKNPYKGWKAPMIDNDFKDDPVIYAFDLSK 318

QY 180 DLQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKMDKQDEORLKEEEDK 239  
Db 319 ELQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKMDKQDEORLKEEEDK 375

QY 240 KKEEEDKDEEDKDEEDKDEEDKDEEDKDEEDKDEEDKDEEDKDEEDKDEEDK 280  
Db 376 PADSDADEDDDDADDTGEDDGSKSDAEDSAEDV----HDEL 415

## RESULT 15

S58170  
calreticulin precursor - maize  
C;Species: Zea mays (maize)  
C;Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 13-Aug-1999

C;Accession: S58170; S49818  
R;Dresselhaus, T.; Hagel, C.; Loerz, H.; Kranz, E.  
submitted to the EMBL Data Library, July 1995  
A;Description: Isolation of a cDNA encoding calreticulin from in vitro zygotes of maize.  
A;Reference number: S58170  
A;Accession: S58170  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-421 <DRE>  
A;Cross-references: EMBL:X89613; NID:g927571; PIDN:CAA61939.1; PID:g927572  
R;Napier, R.M.; Trueman, S.; Henderson, J.; Boyce, J.M.; Hawes, C.R.; Fricker, M.D.; Ven  
submitted to the EMBL Data Library, November 1994  
A;Description: Purification and sequencing of calreticulin from maize and evidence for it  
A;Reference number: S49818  
A;Accession: S49818  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-421 <NAP>  
A;Cross-references: EMBL:Z46772; NID:g577611; PIDN:CAA86728.1; PID:g577612  
C;Genetics:  
A;Gene: CRT1  
C;Superfamily: calreticulin  
C;Keywords: calcium binding  
F;1-25/Domain: signal sequence #status predicted <SIG>  
F;418-421/Region: endoplasmic reticulum retention signal

Query Match 54.3%; Score 844; DB 2; Length 421;  
Best Local Similarity 53.7%; Pred. No. 2.7e-44;  
Matches 154; Conservative 46; Mismatches 71; Indels 16; Gaps 5;

QY 1 GPCTKKVHVFYKGNVILNKIDIRCKDDEFHLYTLVIRPDNTYEVKIDNSQVESGSL 60  
Db 144 GYSTKKVHILTKDGNHLIKKDVPCETDQLTHVTLVIRPDATYSLIDNVEKQTGS 203

QY 61 DWDVFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDW 119  
Db 204 EHWDLFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDERAKIDDPDTSKPEDW 263

QY 120 MDGEWEPVTONPEYKGEWKPQIDNPDPYKGTWVHPEIDNPSPDPSIYVNDNGVLGL 179  
Db 264 EDGEWTAFTIAPNPYKGPWKPKIKNPYKGWKAPMIDNDFKDDPVIYAFDLSK 323

QY 180 DLQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKMDKQDEORLKEEEDK 239  
Db 324 ELQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKMDKQDEORLKEEEDK 379

QY 240 ---KKEEEDKDEEDKDEEDKDEEDKDEEDKDEEDKDEEDKDEEDKDEEDK 280  
Db 380 GGDEDDDDLEDEDEDEKADSKADSDAEDSKSDDEKQ-----HDEL 421

Search completed: October 12, 2004, 15:48:55  
Job time : 18 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 15:46:05 ; Search time 39 Seconds  
(without alignments)  
2265.259 Million cell up

Title: US-09-807-148-9  
 Perfect score: 1553  
 Sequence: 1 GCGTKKVVHIFNYKGNVLI.....EEDKERDEEDVPGQAKDEL 280

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :
SPTRMBL_23.*
1:  sp archaea.*
2:  sp bacteria.*
3:  sp fungi.*
4:  sp human.*
5:  sp invertebrate.*
6:  sp mammal.*
7:  sp mhc.*
8:  sp organelle.*
9:  sp phage.*
10: sp plant.*
11: sp rodent.*
12: sp virus.*
13: sp vertebrate.*
14: sp unclassified.*
15: sp rvirus.*
16: sp bacteriap.*
17: sp archaeap.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282	82.5	418	13	Q7ZXY3 xenopus lae
2	1280.5	82.5	411	13	Q91710 xenopus lae
3	1261.5	81.2	413	13	Q7ZWX8 xenopus lae
4	1241.5	79.9	413	13	Q89364 rana rugosa
5	1198	77.1	423	13	Q7SM3 ictalurus p
6	1184	76.2	318	13	Q9PTX7 lampetra re
7	1164	75.0	405	5	Q36268 aplysia cal
8	1157.5	74.5	417	13	Q80X3 brachydanio
9	1136	73.1	417	13	Q9PUC1 brachydanio
10	1120	72.1	343	13	Q91711 xenopus lae
11	1112	71.6	411	5	Q8MW33 doophilus m
12	1103	71.0	407	5	Q86G72 dermacentor
13	1095.5	70.5	321	13	Q9USG0 epatretus
14	1093.5	70.4	410	5	Q16893 amblyomma a
15	1049.5	67.6	421	5	Q3U6S0 strongyloce
16	1033	66.5	401	5	Q76961 negator ame

Db 139 GPPTKKVHVIFQYKKKNLQINCKDSDFTHTLYTLIVRPNTYEVKIDNSKVSQSGLE 198  
 QY 61 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWKPHEHIPDPDAKPKPDWDEEM 120  
 Db 199 DWDWFLPPKKIKDPDASKPEDWDERPKIDDPDTSKPEDWKPHEHIPDPDAKPKPDWDEEM 258  
 QY 121 DGEWEPVQNPYKGEWKPRQIDNPDKYGTWHPHIDNPYSPDPSIYAYDNFVGLGLD 180  
 Db 259 DGEWEPVQNPYKGEWKPRQIDNPDKYGTWHPHIDNPYSPDPSIYAYDNFVGLGLD 318  
 QY 181 LMQVKSFTIFDNLITNDEAYAEFGNETGWTKAAEKQMKQDEEQLKEEEDKKRK 240  
 Db 319 LMQVKSFTIFDNLITNDEKHAEEYGNETGWTKAAEKQMKQDEEQLKEEEDKKRK 378  
 QY 241 EEE--EAEDKEDDEDKEDEEDKEDEEDVPGQAKDEL 280  
 Db 379 EEPQEEDEDDDDDEEKEEKEEKEEEDDEETP--LKDEL 418  
 RESULT 2  
 Q91710  
 ID Q91710 PRELIMINARY; PRT; 411 AA.  
 AC Q91710;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DE Calreticulin precursor (Fragment).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=CNS;  
 RA Traves S, Zorzato F., Pozzan T.;  
 RT "Identification of calreticulin isoform in the CNS."  
 RL Biochem. J. 0:0-0(0).  
 DR EMBL; X67597; CAA47866.1; -.  
 DR PIR; S29129; S29129.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR003033; Calreticulin.  
 DR InterPro; IPR008985; ConA-like lec\_gi.  
 DR InterPro; IPR00886; ER\_target\_S.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR PRODOM; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 2.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 KW Signal.  
 FT NON TER 1 1  
 FT SIGNAL <1 12 POTENTIAL.  
 FT CHAIN 13 411 CALRETICULIN.  
 SQ SEQUENCE 411 AA; 49344 MW; 8914665008BFA CRC64;  
 Query Match 82.5%; Score 1280.5; DB 13; Length 411;  
 Best Local Similarity 80.4%; Pred. No. 8.8e-79;  
 Matches 225; Conservative 31; Mismatches 22; Indels 3; Gaps 2;  
 QY 133 GPPTKKVHVIFQYKKKNLQINCKDSDFTHTLYTLIVRPNTYEVKIDNSKVSQSGLE 60  
 Db 199 DWDWFLPPKKIKDPDASKPEDWDERPKIDDPDTSKPEDWKPHEHIPDPDAKPKPDWDEEM 252  
 QY 121 DGEWEPVQNPYKGEWKPRQIDNPDKYGTWHPHIDNPYSPDPSIYAYDNFVGLGLD 180  
 Db 253 DGEWEPVQNPYKGEWKPRQIDNPDKYGTWHPHIDNPYSPDPSIYAYDNFVGLGLD 312

QY 181 LMQVKSFTIFDNLITNDEAYAEFGNETGWTKAAEKQMKQDEEQLKEEEDKKRK 240  
 Db 313 LMQVKSFTIFDNLITNDEKHAEEYGNETGWTKAAEKQMKQDEEQLKEEEDKKRK 372  
 QY 241 EEE--EAEDKEDDEDKEDEEDKEDEEDVPGQAKDEL 280  
 Db 373 EEPQEEDEDDDDDEEKEEKEEKEEEDDEETP--LKDEL 411  
 RESULT 3  
 Q7ZWU8  
 ID Q7ZWU8 PRELIMINARY; PRT; 413 AA.  
 AC Q7ZWU8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DE Similar to calreticulin.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC046699; AAH46699.1; -.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR InterPro; IPR008985; ConA-like lec\_gi.  
 DR InterPro; IPR00886; ER\_target\_S.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR PRODOM; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 SQ SEQUENCE 413 AA; 48521 MW; DE002F8F0523772B CRC64;

Query Match 81.2%; Score 1261.5; DB 13; Length 413;  
 Best Local Similarity 78.9%; Pred. No. 1.7e-77;  
 Matches 221; Conservative 29; Mismatches 25; Indels 5; Gaps 1;  
 QY 139 GPPTKKVHVIFQYKKKNLQINCKDSDFTHTLYTLIVRPNTYEVKIDNSKVSQSGLE 198  
 Db 61 DWDWFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWKPHEHIPDPDAKPKPDWDEEM 120  
 Db 199 EDWDFLPFKIKDPDASKPEDWDERPKIDDPDTSKPEDWKPHEHIPDPDAKPKPDWDEEM 258  
 QY 121 DGEWEPVQNPYKGEWKPRQIDNPDKYGTWHPHIDNPYSPDPSIYAYDNFVGLGLD 180  
 Db 259 DGEWEPVQNPYKGEWKPRQIDNPDKYGTWHPHIDNPYSPDPSIYAYDNFVGLGLD 318  
 QY 181 LMQVKSFTIFDNLITNDEAYAEFGNETGWTKAAEKQMKQDEEQLKEEEDKKRK 240  
 Db 319 LMQVKSFTIFDNLITNDEKHAEEYGNETGWTKAAEKQMKQDEEQLKEEEDKKRK 378  
 QY 241 EEE--EAEDKEDDEDKEDEEDKEDEEDVPGQAKDEL 280  
 Db 379 EEPQEEDEDDDDDEEKEEKEEKEEEDDEETP--LKDEL 413

RESULT 4  
 Q98984  
 ID Q98984 PRELIMINARY; PRT; 419 AA.  
 AC Q98984;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)

OC Ictaluridae: Ictalurus.  
OX NCBI\_TaxID=7998;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McConnell T.J., Pitzer J.E.;  
RT "Calreticulin in the Channel Catfish, Ictalurus punctatus."  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY342298; AAQ19852.1; -  
SQ SEQUENCE 423 AA; 49186 MW; ED7CBA5C1FBC07EB CRC64;  
  
Query Match 77.1%; Score 1198; DB 13; Length 423;  
Best Local Similarity 73.7%; Pred. No. 3.3e-73;  
Matches 213; Conservative 36; Mismatches 28; Indels 12; Gaps 3;  
  
QY 1 GPGTKKVVHVFYNYKGNVLINKIRKDKDEFTHLYTLVRPNTYEVKIDNSQVSGSLE 60  
DB 138 GYSTKKVHVFYNYKGNHLLKDKDEFTHLYTLVRPNTYEVKIDNSQVSGSLE 197  
QY 61 DWDWFLPPKKIKDPASKPEDWDERAKIDDDPTDTPEDWDEKPEHIPPDAKPPEDWDEEM 120  
DB 198 EDWDFLPPKKIKDPASKPEDWDERAKIDDDPTDTPEDWDEKPEHIPPDAKPPEDWDEEM 257  
QY 121 DGWEPPVITNPYKGEWKPRQIDNPDKYKGTWHPHPEIDNPYSPPSIYAYDNFVGLGLD 180  
DB 258 DGWEPPVITNPYKGEWKPRQIDNPDKYKGTWHPHPEIDNPYSPPSIYAYDNFVGLGLD 317  
QY 181 LWQVKSGLTFDNLITNDAYAEFGNETGWTKAAEKQKQDEEQRLKEEEDKKRK 240  
DB 318 LWQVKSGLTFDNLITDDVKEAEDFGKGTWGTATKGEKQKQDEEER-KOREEERK 376  
QY 241 EEEAEADKDDKDEDEDEDEK-----EEDDEEDVPGQAKDEL 280  
DB 377 KESEGDGDDDEPEDE 423  
  
RESULT 6  
Q9PTX7 PRELIMINARY; PRT; 318 AA.  
AC Q9PTX7;  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE Calreticulin (fragment).  
OS Lampetra reissneri (Far Eastern brook lamprey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
OC Petromyzontiformes; Petromyzontidae; Lethenteron.  
OX NCBI\_TaxID=7753;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=10594174;  
RA Kuraku S., Hoshiyama D., Katoh K., Suga H., Miyata T.;  
RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded  
genes."  
RL J. Mol. Evol. 49:729-735(1999).  
DR EMBL; AB025328; BAA88481.1; -  
DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
DR InterPro; IPR001580; Calreticulin.  
DR InterPro; IPR009033; Calret. calnex. P.  
DR InterPro; IPR008985; ConA-like lec-gl.  
DR InterPro; IPR000886; ER target S.  
DR Pfam; PF00262; calreticulin; 1.  
DR PRINTS; PS00626; CALRETICULIN.  
DR PRODOM; PD001866; Calreticulin; 1.  
DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
FT NON TER 1  
SQ SEQUENCE 318 AA; 36997 MW; C88102EA1CAC1506 CRC64;  
  
Query Match 76.2%; Score 1184; DB 13; Length 318;  
Best Local Similarity 73.1%; Pred. No. 2.2e-72;  
Matches 207; Conservative 36; Mismatches 36; Indels 4; Gaps 2;

DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE Calreticulin.  
OS Rana rugosa (Wrinkled frog).  
OC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Anura; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.  
OX NCBI\_TaxID=8410;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96234004; PubMed=8654561;  
RA Yamamoto S., Nakamura M.;  
RT "Calnexin: its molecular cloning and expression in the liver of the  
frog, Rana rugosa."  
RL J. Exp. Zool. 275:431-443(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96387817; PubMed=8795287;  
RA Yamamoto S., Kondo Y., Hanada H., Nakamura M.;  
RT "Strong expression of the calreticulin gene in the liver of Rana  
rugosa tadpoles, but not adult frogs."  
RL J. Exp. Zool. 275:431-443(1996).  
DR EMBL; D78589; BAA11425.1; -  
DR PIR; S71343; S71343.  
DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
DR InterPro; IPR001580; Calreticulin.  
DR InterPro; IPR009033; Calret. calnex. P.  
DR InterPro; IPR008985; ConA-like lec-gl.  
DR InterPro; IPR000886; ER target S.  
DR Pfam; PF00262; calreticulin; 1.  
DR PRINTS; PS00626; CALRETICULIN.  
DR PRODOM; PD001866; Calreticulin; 1.  
DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
SQ SEQUENCE 419 AA; 48658 MW; 2C857036769673BF CRC64;  
  
Query Match 79.9%; Score 1241.5; DB 13; Length 419;  
Best Local Similarity 76.9%; Pred. No. 3.8e-76;  
Matches 216; Conservative 34; Mismatches 30; Indels 1; Gaps 1;  
  
QY 1 GPGTKKVVHVFYNYKGNVLINKIRKDKDEFTHLYTLVRPNTYEVKIDNSQVSGSLE 60  
DB 139 GPTTKKVVHVFYNYKGNVLINKIRKDKDEFTHLYTLVRPNTYEVKIDNSQVSGSLE 198  
QY 61 DWDWFLPPKKIKDPASKPEDWDERAKIDDDPTDTPEDWDEKPEHIPPDAKPPEDWDEEM 120  
DB 199 EDWDFLPPKKIKDPASKPEDWDERAKIDDDPTDTPEDWDEKPEHIPPDAKPPEDWDEEM 258  
QY 121 DGWEPPVITNPYKGEWKPRQIDNPDKYKGTWHPHPEIDNPYSPPSIYAYDNFVGLGLD 180  
DB 259 DGWEPPVITNPYKGEWKPRQIDNPDKYKGTWHPHPEIDNPYSPPSIYAYDNFVGLGLD 318  
QY 181 LWQVKSGLTFDNLITNDAYAEFGNETGWTKAAEKQKQDEEQRLKEEEDKKRK 240  
DB 319 LWQVKSGLTFDNLITDDKFAEHAATKGTWGTATKGEKQKQDEEERKQDEEERK 378  
QY 241 EEEAEADKDDKDEDEDEDEK-----EEDDEEDVPGQAKDEL 280  
DB 379 EOEPAEASDDDDDDDDDE 419  
  
RESULT 5  
Q7SZM3 PRELIMINARY; PRT; 423 AA.  
AC Q7SZM3;  
DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE ER-resident chaperone calreticulin.  
OS Ictalurus punctatus (Channel catfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;

QY 1 GPGTKKHVIFNYKGNVLINKIRCKDEFTHTLYTLIVRPDNTYEVKIDNSOVESGSL 60  
 DB 37 GPGTKKHVIFNYKGNVLINKIRCKDEFTHTLYTLIVRPDNTYEVKIDNKVESGSL 96  
 QY 61 DDWDFLPPKKIKDPDASKPDWDERAKIDDPDTSKPEDWDKPEHIIPDPAKKPDWDEM 120  
 DB 97 EDWDLIPAKKIKDPEAKPEDWEDQAKIDDPDVKPDEWKEPEHIIPDPAKKPDWDEM 156  
 QY 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPYSPDSIYVDFNFGVLGLD 180  
 DB 157 DGEWEPPIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPYSPDSIYVDFNFGVLGLD 216  
 QY 181 LQVKSGETIFDNLITNDAYAEFGNETGWTKAAEKQMKQKQDEBQRLKEEEDKKRK 240  
 DB 217 LQVKSGETIFDNLITNDAYAEFGNETGWTKAAEKQMKQKQDEBQRLKEEEDKKRK 276  
 QY 241 EEEAE--DKEDDE 280  
 DB 277 EPEDKDDGDDDDDDDAKAKDDPEPDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD 318

## RESULT 7

Q26268 PRELIMINARY; PRT; 405 AA.  
 AC Q26268  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Calreticulin.  
 GN CALRETICULIN.  
 OS Aplysia californica (California sea hare).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidea;  
 OC Aplysioidae; Aplysiidae; Aplysia.  
 OC NCBI\_TaxID=6500;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93098937; PubMed=1463504;  
 RX Kennedy T.E., Kuhl D., Barzilai A., Sweatt J.D., Kandel E.R.;  
 RT "Long-term sensitization training in Aplysia leads to an increase in  
 calreticulin, a major presynaptic calcium-binding protein.";  
 RL Neuron 9:1013-1024 (1992).  
 DR EMBL; S51239; AAB24569.1; --  
 DR PIR; JH0795; JH0795.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR Pfam; PF00626; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR ProDom; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_REPEAT; 3.  
 DR PROSITE; PS00014; ER TARGET; 1.  
 DR PROSITE; PS00014; ER TARGET; 1.  
 SQ SEQUENCE 405 AA; 46738 MW; 14CAA201840D1D69 CRC64;

Query Match 75.0%; Score 1164; DB 5; Length 405;  
 Best Local Similarity 72.9%; Pred. No. 6.3e-71;  
 Matches 204; Conservative 29; Mismatches 39; Indels 8; Gaps 2;

QY 1 GPGTKKHVIFNYKGNVLINKIRCKDEFTHTLYTLIVRPDNTYEVKIDNSOVESGSL 60  
 DB 134 GPGTKKHVIFNYKGNVLINKIRCKDEFTHTLYTLIVRPDNTYEVKIDNKVESGSL 193  
 QY 61 DDWDFLPPKKIKDPDASKPDWDERAKIDDPDTSKPEDWDKPEHIIPDPAKKPDWDEM 120  
 DB 194 ADWDFLPAKTIIPDPAKKPDWDERAKIDDPDTSKPEDWDKPEHIIPDPAKKPDWDEM 253  
 QY 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPYSPDSIYVDFNFGVLGLD 180  
 DB 254 DGEWEPPIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPYSPDSIYVDFNFGVLGLD 313

QY 181 LQVKSGETIFDNLITNDAYAEFGNETGWTKAAEKQMKQKQDEBQRLKEEEDKKRK 240  
 DB 314 LQVKSGETIFDNLITNDAYAEFGNETGWTKAAEKQMKQKQDEBQRLKEEEDKKRK 373  
 QY 241 EEEAEKDE 280  
 DB 374 EEEA-----NKD-DEEEAEDEEEEDDDDAAPKDEL 405

## RESULT 8

Q802X3 PRELIMINARY; PRT; 417 AA.  
 AC Q802X3  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to calreticulin.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC046906; AAH46906.1; --  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR InterPro; IPR009033; Calreticulin.  
 DR Pfam; PF00626; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR ProDom; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_REPEAT; 3.  
 DR PROSITE; PS00014; ER TARGET; 1.  
 DR PROSITE; PS00014; ER TARGET; 1.  
 SQ SEQUENCE 417 AA; 48259 MW; F7D11614F4319586 CRC64;

Query Match 74.5%; Score 1157.5; DB 13; Length 417;  
 Best Local Similarity 72.1%; Pred. No. 1.8e-70;  
 Matches 202; Conservative 37; Mismatches 40; Indels 1; Gaps 1;

QY 1 GPGTKKHVIFNYKGNVLINKIRCKDEFTHTLYTLIVRPDNTYEVKIDNSOVESGSL 60  
 DB 139 GYSTKXHVIFNYKGNVLINKIRCKDEFTHTLYTLIVRPDNTYEVKIDNKVESGSL 198  
 QY 61 DDWDFLPPKKIKDPDASKPDWDERAKIDDPDTSKPEDWDKPEHIIPDPAKKPDWDEM 120  
 DB 199 EDWDFLPPKKIKDPEAKPEDWDERAKIDDPDTSKPEDWDKPEHIIPDPAKKPDWDEM 258  
 QY 121 DGEWEPPIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPYSPDSIYVDFNFGVLGLD 180  
 DB 259 DGEWEPPIQNPYKGEWKPRQIDNPDKYGTWHPHPEIDNPYSPDSIYVDFNFGVLGLD 318  
 QY 181 LQVKSGETIFDNLITNDAYAEFGNETGWTKAAEKQMKQKQDEBQRLKEEEDKKRK 240  
 DB 319 LQVKSGETIFDNLITNDAYAEFGNETGWTKAAEKQMKQKQDEBQRLKEEEDKKRK 377  
 QY 241 EEEAEKDE 280  
 DB 378 NTEAEEDE 417

## RESULT 9

Q9PUC1 PRELIMINARY; PRT; 417 AA.  
 ID Q9PUC1  
 AC Q9PUC1  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 Calreticulin.  
 GN CALR.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20190113; PubMed=10660676;  
 RA Rubinstein A.L., Lee D., Luo R., Henion P.D., Halpern M.E.;  
 RT "Genes dependent on zebrafish cyclops function identified by AFLP  
 RT differential gene expression screen."  
 RL Genes 26:86-97(2000).  
 RL EMBL; AF195882; AAF13700.1; --  
 DR ZFIN; ZDB-GENE-000208-17; calr.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR009033; Calret\_calmex\_P.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR PRODOM; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 SQ SEQUENCE 417 AA; 48723 MW; 2000CSB4004699B6 CRC64;  
 Query Match 73.1%; Score 1136; DB 13; Length 417;  
 Best Local Similarity 70.5%; Pred. No. 5e-69; Indels 2; Gaps 2;  
 Matches 198; Conservative 42; Mismatches 39;  
 QY 1 GPGTKKHVIFNYKGNVLINRKCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 60  
 DB 138 GYSTKKHVIFNYKGNLKKKDKDDELTHLTLILRPDQYEVKIDNEKVESGSLE 197  
 QY 61 DWDWFLPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDEPHIPDPDAKKPDWDEEM 120  
 DB 198 EDWDFLPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDEPHIPDPDAKKPDWDEEM 257  
 QY 121 DGEWEPPIQNPYKGEWKPQIDNPDKYKWIHPDINPEYSPDPPIYAYDNFGVLGLD 180  
 DB 258 DGEWEPPIQNPYKGEWKPQIDNPDKYKWIHPDINPEYSPDPPIYAYDNFGVLGLD 317  
 QY 181 LWQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKQDEORLKEEBEDKKRK 240  
 DB 318 LWQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKQDEORLKEEBEDKKRK 376  
 QY 241 EEEAEKDDK-DE 280  
 DB 377 DDNEDE 417  
 RESULT 10  
 Q91711 PRELIMINARY; PRT; 343 AA.  
 ID Q91711  
 AC Q91711  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Calreticulin (Fragment).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=CNS;  
 RA Treves S., Zorzato F., Pozzan T.;

"Identification of calreticulin isoform in the CNS."  
 RL Biochem. J. 0:0-0(0).  
 DR EMBL; X67598; CAA47867.1; --  
 DR FIR; S29130; S29130.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR009033; Calret\_calmex\_P.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR PRODOM; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 343 AA; 40105 MW; 3E7DDAP3B91DE1 CRC64;  
 Query Match 72.1%; Score 1120; DB 13; Length 343;  
 Best Local Similarity 83.8%; Pred. No. 4.9e-68;  
 Matches 192; Conservative 21; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 GPGTKKHVIFNYKGNVLINRKCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 60  
 DB 110 GPGTKKHVIFNYKGNLKKKDKDDELTHLTLILRPDQYEVKIDNEKVESGSLE 169  
 QY 61 DWDWFLPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDEPHIPDPDAKKPDWDEEM 120  
 DB 170 EDWDFLPKKIKDPDASKPEDWDERAKIDDDTDSKPEDWDEPHIPDPDAKKPDWDEEM 229  
 QY 121 DGEWEPPIQNPYKGEWKPQIDNPDKYKWIHPDINPEYSPDPPIYAYDNFGVLGLD 180  
 DB 230 DGEWEPPIQNPYKGEWKPQIDNPDKYKWIHPDINPEYSPDPPIYAYDNFGVLGLD 289  
 QY 181 LWQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKQDEORLKEEBEDKKRK 229  
 DB 290 LWQVKSGLTFDNLITNDAYAEFGNETGWVTKAAEKQKQDEORLKEEBEDKKRK 338  
 RESULT 11  
 Q8MW23 PRELIMINARY; PRT; 411 AA.  
 ID Q8MW23  
 AC Q8MW23; 2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Calreticulin.  
 OS Boophilus microplus (Cattle tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Boophilus.  
 OX NCBI\_TaxID=6941;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ferreira C.A.S., Vaz I.S. Jr., Silva S.S., Haag K.L., Valenzuela J.G.,  
 RA Masuda A.;  
 RT "Molecular cloning and partial characterization of a putative  
 RT calreticulin from the cattle tick Boophilus microplus (Acari;  
 RT Ixodidae)."  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF420211; AAN03709.1; --  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR009033; Calret\_calmex\_P.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR PRODOM; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 SQ SEQUENCE 411 AA; 47667 MW; 5FB184E00170A00A CRC64;



Search completed: October 12, 2004, 15:47:23  
Job time : 41 secs

## RESULT 15

Q9U6S0  
ID Q9U6S0  
PRELIMINARY:  
PRT: 421 AA.

Q906S0;  
AC  
DT  
01-MAY-2000 (TREMBl.re) 13: Created)

Q906S0;  
AC  
DT  
01-MAY-2000 (TREMBl.re) 13: Created)

DT	01-MAY-2000	(TREM)Rel. 13,	Last sec
DT	01 MAY 2000	(TREM)rel 13	Last sec

DT	01-MAY-2000	(Tremblrel. 13, Last sequence update)
ET	01-OCT-2002	(Tremblrel. 25, Last annotation update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 15:46:05 ; Search time 13 Seconds  
(without alignments)  
1121.511 Million cell updates/sec

Title: US-09-807-148-9

Perfect score: 1553  
Sequence: 1 GPCTKXVHVIFNYKGNVLI.....EDXDEDEEDVFGQAKDEL 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553	100.0	417	CRTC_HUMAN	P27797 homo sapien
2	1490	95.9	417	CRTC CRIGR	Q8K3h7 cricetus
3	1487.5	95.8	416	CRTC MOUSE	P14211 mus musculus
4	1476.5	95.1	416	CRTC RAT	P18418 rattus norv
5	1457.5	93.9	418	CRTC RABIT	P15253 oryctolagus
6	1442	92.9	417	CRTL BOVIN	P52193 bos taurus
7	1432	92.2	421	CRTL BOVIN	P42918 bos taurus
8	1031.5	66.4	406	CRTC DROME	P29413 drosophila
9	991.5	63.8	395	CRTC CAEEL	P27798 caenorhabdi
10	944.5	60.8	388	RALI ONCVO	P11012 onchocerca
11	856.5	55.2	415	CRTC RICC	P93508 ricinus com
12	837.5	53.9	420	CRTC CHLSE	Q98td3 chlamydomon
13	837.5	53.9	420	CRTC MAIZE	Q98p22 zea mays (m
14	824	53.1	424	CRTL ARATH	Q38858 arabidopsis
15	821.5	52.9	421	CRTC PRUAR	Q9xf98 prunus arme
16	818.5	52.7	416	CRTC NICPL	Q40401 nicotiana p
17	817.5	52.6	393	CRTC SCHMA	Q06814 schistosoma
18	814.5	52.4	416	CRTC BETVU	O81919 beta vulgar
19	810.5	52.2	425	CRTL ARATH	O04151 arabidopsis
20	795	51.2	401	CRTC EUGGR	Q9zny3 euglena gra
21	790.5	50.9	416	CRTC BERST	Q92ppl berberis st
22	779	50.2	424	CRTC ORYSA	Q9sly8 oryza sativ
23	764	49.2	424	CRTC DICDI	Q23858 dictyosteli
24	699.5	45.0	424	CRTL ARATH	O04153 arabidopsis
25	633	40.8	591	CALX_MOUSE	P35564 mus musculu
26	629	40.5	591	CALX_MOUSE	P35565 rattus norv
27	628	40.4	592	CALX_MOUSE	P27824 homo sapien
28	624	40.2	593	CALX CANFA	P24643 canis famil
29	577.5	37.2	610	CALG_HUMAN	O14967 homo sapien
30	562.5	36.2	384	CRTL_HUMAN	Q96112 homo sapien
31	554	35.7	611	CALG_MOUSE	P52194 mus musculu
32	521	34.2	546	CALX_SOYEN	Q39817 glycine max
33	527	33.9	380	CRTL_MOUSE	Q949q6 mus musculu

ALIGNMENTS

RESULT 1

ID	CRTC_HUMAN	STANDARD;	PRT;	417 AA.
AC	P27797			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERP60).			
GN	CALR OR CRTC.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=92013129; PubMed=1919005;			
RA	Rokeach L.A., Hasselby J.A., Meilof J.F., Smeenk R.J., Unnasch T.R.,			
RA	Greene B.M., Hoch S.O.;			
RT	"Characterization of the autoantigen calreticulin.";			
RL	J. Immunol. 147:3031-3039(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=90237213; PubMed=2332496;			
RA	McCaulliffe D.P., Lux F.A., Lieu T.S., Sanz I., Hanke J., Newkirk M.M.,			
RA	Bachinski J.L., Itch Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,			
RA	Capra J.D.;			
RT	"Molecular cloning, expression, and chromosome 19 localization of a			
RT	human Ro/SS-A autoantigen.";			
RL	J. Clin. Invest. 85:1379-1391(1990).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=92129342; PubMed=1733953;			
RA	McCaulliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;			
RA	"The 5'-flanking region of the human calreticulin gene shares			
RT	homology with the human GRP78, GRP94, and protein disulfide isomerase			
RT	promoters.";			
RL	J. Biol. Chem. 267:2557-2562(1992).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RP	Liu J., Peng X., Yuan J., Qiang B.;			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RP	Lamerdin J.E., McCready P.M., Stilwagen S., Ramirez M., Carrano A.;			
RT	"Characterization by genomic sequence analysis of a gene-rich 111 kb			
RT	region of 19p13.2 containing the human DNA repair gene, RAD23A.";			
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=Eye, Pancreas, and Skin;			
RC	MEDLINE=2338825; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			

34	501.5	32.3	530	1	CAX1_ARATH	P29402 arabidopsis
35	493	31.7	551	1	CALX_PEA	O82709 pisum sativ
36	477.5	30.7	532	1	CAX2_ARATH	Q38798 arabidopsis
37	472.5	30.4	560	1	CALX_SCHPO	P36581 schizosach
38	472.5	30.4	619	1	CALX CAEEL	P34652 caenorhabdi
39	451.5	29.1	540	1	CALX HELTU	Q39994 helianthus
40	299.5	19.3	502	1	CALX YEAST	P27825 saccharomyc
41	184.5	11.9	713	1	NUCL_MESAU	P08199 mesocricetu
42	184	11.8	411	1	MP62_LYTP1	P91753 lytechinus
43	183.5	11.8	743	1	ABRA_PLAFC	P22620 plasmodium
44	177.5	11.4	706	1	NUCL_MOUSE	P09405 mus musculu
45	176	11.3	694	1	NUCL_CHICK	P15771 gallus gall

DR	EMBL; M84739; AAA51916.1; -.
DR	EMBL; M32294; AAA36582.1; -.
DR	EMBL; AY047586; AAL13126.1; -.
DR	EMBL; AD000092; AAB51176.1; -.
DR	EMBL; BC002500; AAH02500.1; -.
DR	EMBL; BC007911; AAH07911.1; -.
DR	EMBL; BC020433; AAH20433.1; -.
DR	PIR; A42330; A37047.
DR	PDB; 2CLR; 3I-MAR-95.
DR	SWISS-2DPAGE; P27797; HUMAN.
DR	Aarhus/Ghent-2DPAGE; 9401; IEF.
DR	HSC-2DPAGE; P27797; HUMAN.
DR	PHCI-2DPAGE; P27797; -.
DR	PMAA-2DPAGE; P27797; -.
DR	Siena-2DPAGE; P27797; -.
DR	Genew; HGNC:1455; CALR.
DR	MM; 109091; -.
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
DR	InterPro; IPRO09033; Calret calnex_P.
DR	InterPro; IPRO01580; Calreticulin.
DR	InterPro; IPRO08985; ConA_like Lec-gl.
DR	InterPro; IPRO00886; ER_target_S.
DR	Fram; PF00262; calreticulin; 1.
DR	PRSF; PRSF002356; Calreticulin; 1.
DR	PRINTS; PR00626; CALRETICULIN.
DR	ProDom; PD001866; Calreticulin; 1.
DR	PROSITE; PS00014; ER_TARGET; 1.
DR	PROSITE; PS00803; CALRETICULIN_1; 1.
DR	PROSITE; PS00804; CALRETICULIN_2; 1.
DR	PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW	Endoplasmic reticulum; Calcium-binding; Repeat; Signal; 3D-structure.
FT	SIGNAL 1 17
FT	CHAIN 18 417 CALRETICULIN.
FT	DOMAIN 18 197 N-DOMAIN.
FT	DOMAIN 198 308 P-DOMAIN.
FT	DOMAIN 309 417 C-DOMAIN.
FT	DOMAIN 191 255 4 X APPROXIMATE REPEATS.
FT	REPEAT 191 202 1-1.
FT	REPEAT 210 221 1-2.
FT	REPEAT 227 238 1-3.
FT	REPEAT 244 255 1-4.
FT	DOMAIN 259 297 3 X APPROXIMATE REPEATS.
FT	REPEAT 259 269 2-1.
FT	REPEAT 273 283 2-2.
FT	REPEAT 287 297 2-3.
FT	DOMAIN 351 408 ASP/GLU/LYS-RICH.
FT	DISULFID 137 163 BY SIMILARITY
FT	SITE 414 417 PREVENT SECRETION FROM ER.
FT	CONFLICT 35 35 MISSING (IN REF. 3).
SQ	SEQUENCE 417 AA; 48141 MW; BC37C3CF1054FB2 CRC64;
Query Match 100.0%; Score 1553; DB 1; Length 417;	
Best Local Similarity 100.0%; Pred. No. 2,3e-84;	
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps	
QY	1 GSGTKKHVIFNYKGNVLINKDIRCKDEFFHLTYLIVRPDNTVEVKIDNSQVESGSL
Db	138 GPGTKKHVIFNYKGNVLINKDIRCKDEFFHLTYLIVRPDNTVEVKIDNSQVESGSL
QY	61 DDWDFLPPKKIKDPASKEPDERAKIDDPTDSKPEDWDKPEHIPDPDAKKPEDWDSE
Db	198 DDWDFLPPKKIKDPASKEPDERAKIDDPTDSKPEDWDKPEHIPDPDAKKPEDWDSE
QY	121 DGEWEPPVIQNPEYKGWKPKQIDNPDKGTWIHPIDNPEYSDPSIAYDNFGVLGLD
Db	258 DGEWEPPVIQNPEYKGWKPKQIDNPDKGTWIHPIDNPEYSDPSIAYDNFGVLGLD
QY	181 LMQVKSgttIFNFILTNDDEAVAEFGNETGWVTAAEKQMCKODEORLKEEDEDKKRK
Db	318 LMQVKSgttIFNFILTNDDEAVAEFGNETGWVTAAEKQMCKODEORLKEEDEDKKRK
QY	241 EEEEAEDKEDDDKDEDEBEDEEDEEDVPQAAXDEL 280





RL Biochem. Biophys. Res. Commun. 186:668-673 (1992).

[7]  
 RP SEQUENCE OF 18-32.  
 RC STRAIN=LEC; TISSUE=Liver;  
 RX MEDLINE=94072621; PubMed=8251535;  
 RA Yokoi T., Nagayama S., Kajiwara R., Kawaguchi Y., Horiuchi R.,  
 RA Kamataki T.;  
 RT "Identification of protein disulfide isomerase and calreticulin as  
 RT autoimmunity antigens in LEC strain of rats.";  
 RL Biochim. Biophys. Acta 1158:339-344 (1993).  
 CC -1- FUNCTION: This protein binds calcium. There are both high and low  
 CC affinity calcium-binding sites.  
 CC -1- SUBUNIT: Monomer (BY similarity).  
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.  
 CC -1- SIMILARITY: Belongs to the calreticulin family.  
 CC -1- CAUTION: Was originally (Ref.2) thought to be D-beta-  
 CC hydroxybutyrate dehydrogenase.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; D78308; BAA11345.1; -;  
 DR EMBL; X53363; CAA37446.1; -;  
 DR EMBL; X13702; CAA31987.1; AUT\_SEQ.  
 DR EMBL; X79327; CAA55890.1; -;  
 DR FIR; JH0819; JH0819.  
 DR PDB; 1HN; 26-FEB-02.  
 DR PDB; 1K91; 12-OCT-02.  
 DR PDB; 1K9C; 12-OCT-02.  
 DR InterPro; IPR009033; Calret calnex\_P.  
 DR InterPro; IPR001880; Calreticulin.  
 DR InterPro; IPR008985; ConA-like lec gl.  
 DR InterPro; IPR008886; ER target\_S.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PIRSF; PIRSF002356; Calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR ProDom; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00014; ER-TARGET; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; 3D-structure.  
 FT SIGNAL 1 17  
 FT CHAIN 18 416  
 FT DOMAIN 18 137  
 FT DOMAIN 198 308  
 FT DOMAIN 309 416  
 FT DOMAIN 191 255  
 FT REPEAT 191 202  
 FT REPEAT 210 221  
 FT REPEAT 227 238  
 FT REPEAT 244 255  
 FT DOMAIN 259 277  
 FT REPEAT 259 269  
 FT REPEAT 273 283  
 FT REPEAT 287 297  
 FT DOMAIN 351 407  
 FT DISULFID 137 163  
 FT SITE 413 416  
 SQ SEQUENCE 416 AA; 47995 MW; 285713CE031A2970 CRC64;

Query Match 95.1%; Score 1476.5; DB 1; Length 416;  
 Best Local Similarity 94.6%; Pred. No. 7e-80;  
 Matches 265; Conservative 10; Mismatches 4; Indels 1; Gaps 1;

OY 1 GPGTKVHVIFNYKGNVLINKIDCKDDETHLYTLVRPDNTYEVKIDNSQVESGLE 60

DB 138 GPGTKVHVIFNYKGNVLINKIDCKDDETHLYTLVRPDNTYEVKIDNSQVESGLE 197

OY 61 DDWDFLPKKIKDPDASKPDWDDEAKIDDDTDSKPDWDKPEHIPDPAKPPDWDDEEM 120  
 DB 198 DDWDFLPKKIKDPDASKPDWDDEAKIDDDTDSKPDWDKPEHIPDPAKPPDWDDEEM 257  
 OY 121 DGEWPPVQIONPEYKGEWKPRQIDNPDKYGTWHPIDNPDPSPSIYAYDNFVGLGLD 180  
 DB 258 DGEWPPVQIONPEYKGEWKPRQIDNPDKYGTWHPIDNPDPSPSIYAYDNFVGLGLD 317  
 OY 181 LQVKSQGITFDNFLTNDDEAVAEFNGETWGTAAEKQMKDKODEEORLKEEEDKKRK 240  
 DB 318 LQVKSQGITFDNFLTNDDEAVAEFNGETWGTAAEKQMKDKODEEORLKEEEDKKRK 377  
 OY 241 EEEAEDEKDEDDKDEDEEDKEEDEEDVPGQAKDEL 280  
 DB 378 EEEAEDEKDEDDKDEDEEDKEEDEED-EDATQAKDEL 416

# RESULT 5

CRTC\_RABIT STANDARD; PRT; 418 AA.  
 AC P15253;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERp60).  
 GN CALR.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=slow-twitch skeletal muscle;  
 RX MEDLINE=90094320; PubMed=2600080;  
 RA Fliegel L., Burns K., MacLennan D.H., Reithmeier R.A.F., Michalak M.;  
 RT "Molecular cloning of the high affinity calcium-binding protein  
 RT (calreticulin) of skeletal muscle sarcoplasmic reticulum.";  
 RL J. Biol. Chem. 264:21522-21528 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=fast-twitch skeletal muscle;  
 RX MEDLINE=91282795; PubMed=2059224;  
 RA Fliegel L., Michalak M.;  
 RT "Fast-twitch and slow-twitch skeletal muscles express the same  
 RT isoform of calreticulin.";  
 RL Biochem. Biophys. Res. Commun. 177:979-984 (1991).  
 RN [3]  
 RP SEQUENCE OF 18-36.  
 RX MEDLINE=9105414; PubMed=2241926;  
 RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,  
 RA MacLennan D.H., Meldolesi J., Pozzan T.;  
 RT "Calreticulin is a candidate for a calsequestrin-like function in  
 RT Ca2(+)-storage compartments (calciosomes) of liver and brain.";  
 RL Biochem. J. 271:473-480 (1990).  
 RN [4]  
 RP SEQUENCE OF 18-46.  
 RX MEDLINE=91201375; PubMed=2016321;  
 RA Milner R.E., Bakes S., Shemanko C., Carpenter M.R., Smillie L.,  
 RA Vance J.E., Opas M., Michalak M.;  
 RT "Calreticulin, and not calsequestrin, is the major calcium binding  
 RT protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic  
 RT reticulum.";  
 RL J. Biol. Chem. 266:7155-7165 (1991).  
 RN [5]  
 RP PARTIAL SEQUENCE.  
 RC TISSUE=Lung;  
 RX MEDLINE=92002038; PubMed=1911780;  
 RA Guan S., Fallick A.M., Williams D.E., Cashman J.R.;  
 RT "Evidence for complex formation between rabbit lung flavin-containing  
 RT monooxygenase and calreticulin.";  
 RL Biochemistry 30:9892-9900 (1991).  
 CC -1- FUNCTION: This protein binds calcium. There are both high and low



```

FT DISULFID 137 163
FT CARBOHYD 179 179
FT SITE 414 417
SQ SEQUENCE 417 AA; 48038 MW; 7B9812C7B54178E9 CRC64;

Query Match 92.9%; Score 1442; DB 1; Length 417;
Best Local Similarity 92.9%; Pred. No. 7.3e-78;
Matches 261; Conservative 11; Mismatches 7; Indels 2; Gaps 2;

QY 1 GPGTKVHVIFNYKGNVLINLKDIRCKDDEFTHLYTLVIRPNTVEYKIDNSQVESGSL 60
DB 138 GPGTKVHVIFNYKGNVLINLKDIRCKDDEFTHLYTLVIRPNTVEYKIDNSQVESGSL 197

QY 61 DWDWFLPPKKIKDPDASKEDWDERAKIDDDPSKPEDWDKPEHIPPDAKKPDWDEEM 120
DB 198 DWDWFLPPKKIKDPDASKEDWDERAKIDDDPSKPEDWDKPEHIPPDAKKPDWDEEM 257

QY 121 DGWEPPVQNPPEYKGEKWPQIDNPDKYGTWHPHIDNPPEYSPDPSTIAYDNFVGLD 180
DB 258 DGWEPPVQNPPEYKGEKWPQIDNPDKYGTWHPHIDNPPEYSPDPSTIAYDNFVGLD 317

QY 181 LQVKSGLTFDNLITNDAYAEFGNETGWYTKAAEKQKQDEEORLKEEEDKKRK 240
DB 318 LQVKSGLTFDNLITNDAYAEFGNETGWYTKAAEKQKQDEEORLKEEEDKKRK 377

QY 241 EEEAEADKDEDDKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
DB 378 EEEAEADKDEDDKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 417

RESULT 7
CRT2_BOVIN STANDARD; PRT; 421 AA.
AC P42918;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Calreticulin, Brain Isoform 2 precursor (CRP55) (Calregulin) (HACBP).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93385184; PubMed=8373827;
RA Liu N., Fine R.E., Johnson R.J.;
RT "Comparison of cDNAs from bovine brain coding for two isoforms of
RT calreticulin."
RL Biochim. Biophys. Acta 1202:70-76 (1993).
CC -!- FUNCTION: This protein binds calcium. There are both high and low
CC affinity calcium-binding sites.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Belongs to the calreticulin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L13462; AAC37307.1; --
CC PIR; S36799; S36799.
CC InterPro; IPR009033; Calret_calmex.P.
CC InterPro; IPR001580; Calreticulin.
CC InterPro; IPR008985; ConA like lec_gl.
CC InterPro; IPR000886; R target_S.
CC Pfam; PF00262; calreticulin_1.
CC PIRSF; PIRSF002356; Calreticulin; 1.

```

```

DR PRINTS; P00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
DR PROSITE; PS00014; ER_TARGET; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal.
FT SIGNAL 1 34
FT CHAIN 35 421
FT DOMAIN 35 201
FT DOMAIN 202 312
FT DOMAIN 313 421
FT DOMAIN 195 259
FT REPEAT 195 206
FT REPEAT 214 225
FT REPEAT 231 242
FT REPEAT 248 259
FT DOMAIN 263 301
FT REPEAT 263 273
FT REPEAT 277 287
FT REPEAT 291 301
FT DOMAIN 366 411
FT DISULFID 141 167
FT CARBOHYD 183 183
FT SITE 418 421
SQ SEQUENCE 421 AA; 48812 MW; 0257E959F71528BC CRC64;

Query Match 92.2%; Score 1432; DB 1; Length 421;
Best Local Similarity 92.2%; Pred. No. 2.8e-77;
Matches 259; Conservative 12; Mismatches 8; Indels 2; Gaps 2;

QY 1 GPGTKVHVIFNYKGNVLINLKDIRCKDDEFTHLYTLVIRPNTVEYKIDNSQVESGSL 60
DB 142 GPGTKVHVIFNYKGNVLINLKDIRCKDDEFTHLYTLVIRPNTVEYKIDNSQVESGSL 201

QY 61 DWDWFLPPKKIKDPDASKEDWDERAKIDDDPSKPEDWDKPEHIPPDAKKPDWDEEM 120
DB 202 DWDWFLPPKKIKDPDASKEDWDERAKIDDDPSKPEDWDKPEHIPPDAKKPDWDEEM 261

QY 121 DGWEPPVQNPPEYKGEKWPQIDNPDKYGTWHPHIDNPPEYSPDPSTIAYDNFVGLD 180
DB 262 DGWEPPVQNPPEYKGEKWPQIDNPDKYGTWHPHIDNPPEYSPDPSTIAYDNFVGLD 321

QY 181 LQVKSGLTFDNLITNDAYAEFGNETGWYTKAAEKQKQDEEORLKEEEDKKRK 240
DB 322 LQVKSGLTFDNLITNDAYAEFGNETGWYTKAAEKQKQDEEORLKEEEDKKRK 381

QY 241 EEEAEADKDEDDKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 280
DB 382 EEEAEADKDEDDKDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 421

```

```

RESULT 8
CRT2_DROME STANDARD; PRT; 406 AA.
AC P29413; Q9VHA3;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Calreticulin precursor (CRP55) (Calregulin) (HACBP).
GN CRC OR C69429.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208374; PubMed=1296819;
RA Smith M.J.;
RT "Nucleotide sequence of a Drosophila melanogaster gene encoding a
RT calreticulin homologue."
RL DNA Seq. 3:247-250 (1992).

```





CC -!- SIMILARITY: Belongs to the calreticulin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; X59589; CA442159.1; -;  
 CC EMBL; AF125963; AAD14746.1; -;  
 CC PIR; S25851; S25851.  
 CC Wormpep; Y38A10A.5; CR21562.  
 CC InterPro; IPR009033; Calret\_calmex\_P.  
 CC InterPro; IPR001580; Calreticulin.  
 CC InterPro; IPR008985; ER\_target\_S.  
 CC InterPro; IPR000886; ER\_target\_S.  
 CC Pfam; PF00262; calreticulin; 1.  
 CC PIRSF; PIRSF002356; Calreticulin; 1.  
 CC PRINTS; PR00626; CALRETICULIN.  
 CC ProDom; PD001866; Calreticulin; 1.  
 CC PROSITE; PS00014; ER\_TARGET; 1.  
 CC PROSITE; PS00803; CALRETICULIN\_1; 1.  
 CC PROSITE; PS00804; CALRETICULIN\_2; 1.  
 CC PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 CC KEGG; Endoplasmic reticulum; Calcium-binding; Repeat; Signal.  
 CC SIGNAL 1 15 POTENTIAL.  
 FT CHAIN 16 395 CALRETICULIN.  
 FT DOMAIN 2 192 N-DOMAIN.  
 FT DOMAIN 193 301 P-DOMAIN.  
 FT DOMAIN 302 395 C-DOMAIN.  
 FT DOMAIN 186 250 4 X APPROXIMATE REPEATS.  
 FT REPEAT 186 197 1-1.  
 FT REPEAT 205 216 1-2.  
 FT REPEAT 222 233 1-3.  
 FT REPEAT 239 250 1-4.  
 FT DOMAIN 254 292 3 X APPROXIMATE REPEATS.  
 FT REPEAT 254 264 2-1.  
 FT REPEAT 268 278 2-2.  
 FT REPEAT 282 292 2-3.  
 FT DOMAIN 332 390 ASP/GLU/LYS-RICH.  
 FT DISULFID 133 158 BY SIMILARITY.  
 FT SITE 392 395 PREVENT SECRETION FROM ER.  
 CC SEQUENCE 395 AA; 45616 MW; 35CA7D2EC1D56B03 CRC64;  
 Query Match 63.8%; Score 991.5; DB 1; Length 395;  
 Best Local Similarity 66.0%; Pred. No. 1.5e-51;  
 Matches 173; Conservative 35; Mismatches 53; Indels 1; Gaps 1;  
 QY 1 GPQTKVHVIFNYKGNVLINCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSLE 60  
 DB 134 GP-TRRVHVLNKGKGLIKKEITCKSDLTHLYTLINSDNTYEVKIDGESAQTSLE 192  
 QY 61 DDWDFLPPKKIKDPDASKEDWDERAKIDDPDTSKPDWDERKPEHIPPDPKAKPEDWDEEM 120  
 DB 193 EDWDLFPKKIKDPDASKEDWDEREYIDDAEDAKPEDWKEPEHIPPDPKAKPEDWDEEM 252  
 QY 121 DGEWEPVQNPYKGEWPKQIDNPDKYGTWHPIDNPESYDPSPYAYDNTFGVLGLD 180  
 DB 253 DGEWEPVQNPYKGEWPKQIKQPAYKGIWHPIDNPESYDPSPYAYDNTFGVLGLD 312  
 QY 181 LWQVKGSTIFDNLINDRAYAEFGNETWGTVAEKQKQDKQDEEORLKEEEDKGRK 240  
 DB 313 LWQVKGSTIFDNLINDRAYAEFGNETWGTVAEKQKQDKQDEEORLKEEEDKGRK 372  
 QY 241 EEEAEADKDEEDKDEDEDEE 262  
 DB 373 EEEAKKDDDEEKEEKEEGHDE 394

RESULT 10

RALI\_ONCVO

ID RALI\_ONCVO STANDARD; PRT; 388 AA.  
 AC P11012;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE RAL-1 protein precursor (41 kDa larval antigen).  
 GN RAL1.  
 OS Onchocerca volvulus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Onchocerca.  
 OX NCBI\_TaxID=6282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94341871; PubMed=7520419;  
 RA Rokeach L.A., Zimmerman P.A., Unnasch T.R.;  
 RT "Epitopes of the Onchocerca volvulus RAL1 antigen, a member of the  
 RT calreticulin family of proteins, recognized by sera from patients  
 RT with onchocerciasis.";  
 RL Infect. Immun. 62:3696-3704 (1994).  
 RN [2]  
 RP SEQUENCE OF 53-388 FROM N.A.  
 RX MEDLINE=8273584; PubMed=2455736;  
 RA Unnasch T.R., Gallin M.Y., Soboslay P.T., Ertmann K.D., Greene B.M.;  
 RT "Isolation and characterization of expression cDNA clones encoding  
 RT antigens of Onchocerca volvulus infective larvae.";  
 RL J. Clin. Invest. 82:262-269 (1988).  
 CC -!- SIMILARITY: Belongs to the calreticulin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; M20565; AAA59056.1; -;  
 DR PIR; A32507; A32507.  
 DR InterPro; IPR009033; Calret\_calmex\_P.  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR008985; ConA\_like lec.gl.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PIRSF; PIRSF002356; Calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN.  
 DR ProDom; PD001866; Calreticulin; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR Calcium-binding; Repeat; Antigen; Signal.  
 KW SIGNAL 1 17 POTENTIAL.  
 FT CHAIN 18 388 RAL-1 PROTEIN.  
 FT DOMAIN 189 253 4 X APPROXIMATE REPEATS.  
 FT REPEAT 189 200 1-1.  
 FT REPEAT 208 219 1-2.  
 FT REPEAT 225 236 1-3.  
 FT REPEAT 242 253 1-4.  
 FT DOMAIN 257 295 3 X APPROXIMATE REPEATS.  
 FT REPEAT 257 267 2-1.  
 FT REPEAT 271 281 2-2.  
 FT REPEAT 285 295 2-3.  
 FT DOMAIN 353 388 ARG(LYS-RICH (BASIC)).  
 FT DISULFID 135 161 BY SIMILARITY.  
 CC SEQUENCE 388 AA; 45298 MW; 9537F298A2D31CD6 CRC64;  
 Query Match 60.8%; Score 944.5; DB 1; Length 388;  
 Best Local Similarity 65.4%; Pred. No. 8.5e-49;  
 Matches 166; Conservative 32; Mismatches 49; Indels 7; Gaps 2;  
 QY 1 GPQTKVHVIFNYKGNVLINCKDDETHLYTLVIRPDNTYEVKIDNSQVSGSLE 60  
 DB 136 GPQTKVHVIFNYKGNVLINCKDDETHLYTLVIRPDNTYEVKIDGESAESELE 195  
 QY 61 DDWDFLPPKKIKDPDASKEDWDERAKIDDPDTSKPDWDERKPEHIPPDPKAKPEDWDEEM 120



FT SITE 417 420 PREVENT SECRETION FROM ER (POTENTIAL).  
 SQ SEQUENCE 420 AA; 47327 MW; DD3BA3AFFBF61C9B CRC64;

Query Match 53.9%; Score 837.5; DB 1; Length 420;  
 Best Local Similarity 54.0%; Pred. No. 1.7e-42;  
 Matches 154; Conservative 43; Mismatches 65; Indels 23; Gaps 4;

QY 1 GPGTKVHVFNFKGNVINKDIRCKDDEFHLYTLIVRPDNTVEVKIDNSQVSGSLE 60  
 DB 141 GYSTRKVVHLYTKGNLYIKDKIAETQOLTHVTVFIRPDATVSIILIDNEKHTGSIYEHM 200  
 QY 61 DMDLFPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDK-PEHIPDPDAKKPEDWDDE 119  
 DB 201 EDWMLPPKTIKPKASKPEDWDEREIAIDPDKPEGWDDIPATIAKDAKKPEDWDE 260  
 QY 120 MDGEWPPVIONPEYKGEWKPRQIDNPDKYGTWHPDIIDNPYSDDPSIYAYDNFGLGL 179  
 DB 261 EDGTWEPFMPINPEYKGEWKAKMKIPAYKGIWADPIDNPYVHDDKLYNFKULKFVG 320  
 QY 180 DLQVKSGLTFDNLITNDDEAYAEFGNETGWYTKAAEKQMKDKQDEBORLKEEEDKKR 239  
 DB 321 ELQVKSGLTFDNLITVDLEAAKPAEDTWGKHDEKXAFDK-----VKKEEDENKA 374  
 QY 240 KE-----BEAEKEDDEDKDE-----EDEDKEDE 268  
 DB 375 KDAPPVPVDAEAEDEDEYEDKEEPESGMSIKIPKEEESGHDE 419

RESULT 13  
 CRT2 MAIZE  
 ID CRT2 MAIZE STANDARD; PRT; 420 AA.  
 AC Q9SP2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Calreticulin precursor.  
 GN CRT.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21862796; PubMed=11874098;  
 RA Wyatt S.E., Tsou P.-L., Robertson D.,  
 RT "Expression of the high capacity calcium-binding domain of  
 RT calreticulin increases bioavailable calcium stores in plants.";  
 RL Transgenic Res. 11:1-10(2002).  
 CC -!- FUNCTION: This protein binds calcium. There are both high and low  
 CC affinity calcium-binding sites (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).  
 CC -!- SIMILARITY: Belongs to the calreticulin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch))  
 CC -----  
 CC EMBL: AF190454; AAF01470.1;  
 CC InterPro: IPR009033; Calret\_calmex.P.  
 CC InterPro: IPR001580; Calreticulin.  
 CC InterPro: IPR008985; ConA like lec\_gl.  
 CC InterPro: IPR000886; ER\_target.S.  
 CC Pfam: PF00262; calreticulin\_1.  
 CC PRF: PRSF002356; Calreticulin; 1.  
 CC PRINTS: PR00626; CALRETICULIN.  
 CC ProDom: PD001866; Calreticulin; 1.  
 CC PROSITE: PS00014; ER\_TARGET; 1.  
 CC PROSITE: PS00803; CALRETICULIN\_1; 1.

DR PROSITE; PS00804; CALRETICULIN\_2; FALSE NEG.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 2.  
 KW Endoplasmic reticulum; Calcium-binding; Repeat; Signal; Glycoprotein.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 420 CALRETICULIN.  
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT SITE 417 420 PREVENT SECRETION FROM ER (POTENTIAL).  
 SQ SEQUENCE 420 AA; 47939 MW; E73B7F43E7494735 CRC64;

Query Match 53.9%; Score 837.5; DB 1; Length 420;  
 Best Local Similarity 52.9%; Pred. No. 1.7e-42;  
 Matches 147; Conservative 49; Mismatches 77; Indels 5; Gaps 2;

QY 4 TKVHVHFNFKGNVINKDIRCKDDEFHLYTLIVRPDNTVEVKIDNSQVSGSLEDDW 63  
 DB 147 TKVHVHILTKDGNHLLKDVQCFQDQLTHVTVFIRPDATVSIILIDNEKHTGSIYEHM 206  
 QY 64 DFLPPKKIKDPDASKPEDWDERAKIDDPDTSKPEDWDK-PEHIPDPDAKKPEDWDEWDG 122  
 DB 207 DILPPKKIKDPDASKPEDWDDKEYIPDPECKKPEGYDDIPKEIPDPAKKPEDWDEEDG 266  
 QY 123 EWEPPVIONPEYKGEWKPRQIDNPDKYGTWHPDIIDNPYSDDPSIYAYDNFGLGLDW 182  
 DB 267 EWTAPTIPNPEYKGPWKQKKIKPNYQGWKAPMIDNPDFKDDPYIYAFDSLIKYGIELM 326  
 QY 183 QVKSGLTFDNLITNDDEAYAEFGNETGWYTKAAEKQMKDKQDEBORLKEEEDKKRKE 242  
 DB 327 QVKSGLTFDNLITDDPALAKTFAEETWGHKEAKAFADEAK----KKEEEDAAGKD 382  
 QY 243 EAEKEDDEDKDE 280  
 DB 383 DEDDDE 420

RESULT 14  
 CRT2 ARATH  
 ID CRT2 ARATH STANDARD; PRT; 424 AA.  
 AC Q38858; O04152; O80486;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Calreticulin 2 precursor.  
 GN CRT2 OR CRTL OR ATIG09210 OR T12M4.8.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Dewar K.,  
 RA Dunn P., Eggu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen M.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti B., Marziani A.,  
 RA Miliuteer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Taiton L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 408:916-920(2000).  
 RN [2]  
 RP SEQUENCE OF 1-174 FROM N.A.



QY 120 MDGEWPPVIONPEYKGEWKPRQIDNPDYKGTWHPIDNPEYSPDPSIYAYDNFGVLGL 179  
Db 263 EDGEWTAPTIPNPEYKGEWKPKKIKNPNFKGKWKAPLIDNPEFKDDPELYVYPNLKYVGI 322  
QY 180 DLQVKSGTIFDNFLITNDAYAEFFGNETGVTKAAEQMKDKQDEORLKEEEDKKR 239  
Db 323 ELQVKSGLFDNILLITDEPEYAKQLABETWKGKQDAKA---AFEELEKQLQEEESKED 379  
QY 240 KEEEAEDKEDDE 280  
Db 380 PVDSDAED-DDNEAEDGEESDSESKPDSTERSAETEAKEHDEL 421

Search completed: October 12, 2004, 15:46:32  
Job time : 14 secs

**This Page Blank (uspto)**